

Customer No. 20350

TOWNSEND and TOWNSEND and CREW LLP  
Two Embarcadero Center, 8<sup>th</sup> Floor  
San Francisco, California 94111-3834  
(415) 576-0200

ASSISTANT COMMISSIONER FOR PATENTS  
BOX PATENT APPLICATION  
Washington, D.C. 20231

Sir:

Transmitted herewith for filing under 37 CFR 1.53(b) is the

- ☐ patent application of  
☐ continuation patent application of  
☐ divisional patent application of  
☒ continuation-in-part patent application of

Attorney Docket No. 18623-013900US

Client Ref No. EPI 0139.00 US

"Express Mail" Label No. EL11555026US

Date of Deposit: July 8, 1999

I hereby certify that this is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above, addressed to:

Assistant Commissioner for Patents  
Washington, D.C. 20231

By: 

Inventor(s)/Applicant Identifier: Alessandro SETTE, John SIDNEY, Scott SOUTHWOOD, Maria A. VITIELLO,

Brian D. LIVINGSTON, Esteban CELIS, Ralph T. KUBO, Howard M. GREY,  
and Robert CHESNUT

For: INDUCING CELLULAR IMMUNE RESPONSES TO HEPATITIS B VIRUS USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS

☒ This application claims priority from each of the following Application Nos./filing dates: 08/820,360, filed March 12, 1997; 60/013,363, filed March 13, 1996; 09/189,702, filed November 10, 1998; 08/205,713, filed March 4, 1994; 08/159,184, filed November 29, 1993; 08/073,205, filed June 4, 1993; and 08/027,146, filed March 5, 1993, the disclosure(s) of which is (are) incorporated by reference.

☐ Please amend this application by adding the following before the first sentence: "This application is a ☐ continuation ☐ continuation-in-part of and claims the benefit of U.S. Application No. 60/\_\_\_\_\_, filed \_\_\_\_\_, the disclosure of which is incorporated by reference."

Enclosed are:

- ☒ 208 page(s) of specification  
☒ 6 page(s) of claims  
☒ 1 page of Abstract  
☒ 2 sheet(s) of ☐ formal ☒ informal drawing(s).

An assignment of the invention to \_\_\_\_\_

A ☐ signed ☐ unsigned Declaration & Power of Attorney

☒ A ☐ signed ☒ unsigned Declaration.

☐ A Power of Attorney.

☐ A verified statement to establish small entity status under 37 CFR 1.9 and 37 CFR 1.27 ☐ is enclosed ☐ was filed in the prior application and small entity status is still proper and desired.

☐ A certified copy of a \_\_\_\_\_ application.

☐ Information Disclosure Statement under 37 CFR 1.97.

☐ A petition to extend time to respond in the parent application.

☐ Notification of change of ☐ power of attorney ☐ correspondence address filed in prior application.

In view of the Unsigned Declaration as filed with this application and pursuant to 37 CFR §1.53(f),  
Applicant requests deferral of the filing fee until submission of the Missing Parts of Application.

DO NOT CHARGE THE FILING FEE AT THIS TIME.

Telephone:  
(415) 576-0200

Facsimile:  
(415) 576-0300

  
Jean M. Lockyer, Reg No.: P-44,879  
Attorneys for Applicant

**PATENT APPLICATION**

**INDUCING CELLULAR IMMUNE RESPONSES TO HEPATITIS B VIRUS  
USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS**

Inventor(s): Alessandro Sette, an Italian citizen, residing at  
5551 Linda Rosa Avenue  
La Jolla, California 92037

John Sidney, a United States citizen, residing at  
4218 Corte de la Siena  
La Jolla, California 92130

Scott Southwood, a United States citizen, residing at  
10679 Strathmore Drive  
Santee, California 92071

Maria A. Vitiello, an Italian citizen, residing at  
7522 High Avenue  
La Jolla, CA 92037

Brian D. Livingston, a United States citizen, residing at  
13555 Chaco Court  
San Diego, California 92129

Esteban Celis, a United States citizen, residing at  
3683 Wright Road S.W.  
Rochester, Minnesota 55902

Ralph T. Kubo, a United States citizen, residing at  
6921 Pear Tree Drive  
Carlsbad, California 92009

Howard M. Grey, a United States citizen, residing at  
1461 Caminito Batea  
La Jolla, California 92037

Robert Chesnut, a United States citizen, residing at  
1473 Kings Cross Drive  
Cardiff-by-the-Sea, California 92007

TOWNSEND and TOWNSEND and CREW LLP  
Two Embarcadero Center, 8<sup>th</sup> Floor  
San Francisco, California 94111-3834  
(415) 576-0200

PATENT

Attorney Docket No.: 018623-013900US

5     **INDUCING CELLULAR IMMUNE RESPONSES TO HEPATITIS B VIRUS USING  
PEPTIDE AND NUCLEIC ACID COMPOSITIONS**

**CROSS-REFERENCES TO RELATED APPLICATIONS**

          This application is a Continuation-In-Part ("CIP") of U.S.S.N. 08/820,360 filed  
3/12/97, which claims the benefit of U.S. Provisional Application No. 60/013,363 filed  
10    March 13, 1996 and now abandoned. The present application is also a CIP of U.S.S.N.  
09/189,702 filed 11/10/98, which is a CIP of U.S.S.N. 08/205,713 filed 3/4/94, which is a  
CIP of 08/159,184 filed 11/29/93 and now abandoned, which is a CIP of 08/073,205 filed  
6/4/93 and now abandoned, which is a CIP of 08/027,146 filed 3/5/93 and now abandoned.  
The present application is also related to U.S.S.N. 08/197,484, U.S.S.N. 08/464,234,  
15    U.S.S.N. 08/464,496, U.S.S.N. 08/464,031, abandoned U.S.S.N. 08/464,433, and U.S.S.N.  
08/461,603, which is a continuation of abandoned U.S.S.N. 07/935,811, which is a CIP of  
abandoned U.S.S.N. 07/874,491, which is a CIP of abandoned U.S.S.N. 07/827,682, which  
is a CIP of abandoned 07/749,568. The present application is also related to U.S. Patent  
Application entitled "Peptides and Methods for Creating Synthetic Peptides with Modulated  
20    Binding Affinity for HLA Molecules", Attorney Docket No. 018623-009520, filed 1/6/99,  
which is a CIP of U.S.S.N. 08/815,396, which is a CIP of abandoned U.S.S.N. 60/013,113.  
Furthermore, the present application is related to U.S.S.N. 09/017,735, which is a CIP of  
abandoned U.S.S.N. 08/589,108; U.S.S.N. 08/753,622, U.S.S.N. 08/822,382, abandoned  
U.S.S.N. 60/013,980, U.S.S.N. 08/454,033, U.S.S.N. 09/116,424, U.S.S.N. 08/205,713, and  
25    U.S.S.N. 08/349,177, which is a CIP of abandoned U.S.S.N. 08/159,184, which is a CIP of  
abandoned U.S.S.N. 08/073,205, which is a CIP of abandoned U.S.S.N. 08/027,146. The  
present application is also related to U.S.S.N. 09/017,524, U.S.S.N. 08/821,739, abandoned  
U.S.S.N. 60/013,833, U.S.S.N. 08/758,409, U.S.S.N. 08/589,107, U.S.S.N. 08/451,913,  
U.S.S.N. 08/186,266, U.S.S.N. 09/116,061, and U.S.S.N. 08/347,610, which is a CIP of  
30    U.S.S.N. 08/159,339, which is a CIP of abandoned U.S.S.N. 08/103,396, which is a CIP of  
abandoned U.S.S.N. 08/027,746, which is a CIP of abandoned U.S.S.N. 07/926,666. The  
present application is also related to U.S.S.N. 09/017,743, U.S.S.N. 08/753,615; U.S.S.N.  
08/590,298, U.S.S.N. 09/115,400, and U.S.S.N. 08/452,843, which is a CIP of U.S.S.N.

08/344,824, which is a CIP of abandoned U.S.S.N. 08/278,634. The present application is also related to provisional U.S.S.N. 60/087,192 and U.S.S.N. 09/009,953, which is a CIP of abandoned U.S.S.N. 60/036,713 and abandoned U.S.S.N. 60/037,432. In addition, the present application is related to U.S.S.N. 09/098,584 and to Provisional U.S.S.N.

5 60/117,486. All of the above applications are incorporated herein by reference.

### **FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT**

This invention was funded, in part, by the United States government under grants with the National Institutes of Health. The U.S. government has certain rights in this  
10 invention.



## INDEX

	I.	Background of the Invention
	II.	Summary of the Invention
	III.	Brief Description of the Figures
5	IV.	Detailed Description of the Invention
	A.	Definitions
	B.	Stimulation of CTL and HTL responses against HBV
	C.	Binding Affinity of Peptide Epitopes for HLA Molecules
	D.	Peptide Epitope Binding Motifs and Supermotifs
10	1.	HLA-A1 supermotif
	2.	HLA-A2 supermotif
	3.	HLA-A3 supermotif
	4.	HLA-A24 supermotif
	5.	HLA-B7 supermotif
15	6.	HLA-B27 supermotif
	7.	HLA-B44 supermotif
	8.	HLA-B58 supermotif
	9.	HLA-B62 supermotif
	10.	HLA-A1 motif
20	11.	HLA-A2.1 motif
	12.	HLA-A3 motif
	13.	HLA-A11 motif
	14.	HLA-A24 motif
	15.	HLA-DR-1-4-7 supermotif
25	16.	HLA-DR3 motifs
	E.	Enhancing Population Coverage of the Vaccine
	F.	Immune Response Stimulating Peptide Analogs
	G.	Computer Screening of Protein Sequences from Disease-Related Antigens for Supermotif or Motif Containing Peptides
30	H.	Preparation of Peptide Epitopes
	I.	Assays to Detect T-Cell Responses
	J.	Use of Peptide Epitopes for Evaluating Immune Responses
	K.	Vaccine Compositions
	1.	Minigene Vaccines

## 2. Combinations of CTL Peptides with Helper Peptides

## L. Administration of Vaccines for Therapeutic or Prophylactic Purposes

## M. Kits

## V. Examples

## 5 VI. Claims

## VII. Abstract

## I. BACKGROUND OF THE INVENTION

Chronic infection by hepatitis B virus (HBV) affects at least 5% of the world's population and is a major cause of cirrhosis and hepatocellular carcinoma (Hoofnagle, J., *N. Engl. J. Med.* 323:337, 1990; Fields, B. and Knipe, D., In: *Fields Virology* 2:2137, 1990). The World Health Organization lists hepatitis B as a leading cause of death worldwide, close behind chronic pulmonary disease, and more prevalent than AIDS. Chronic HBV infection can range from an asymptomatic carrier state to continuous hepatocellular necrosis and inflammation, and can lead to hepatocellular carcinoma.

The immune response to HBV is believed to play an important role in controlling hepatitis B infection. A variety of humoral and cellular responses to different regions of HBV including the nucleocapsid core, polymerase, and surface antigens have been identified. T cell-mediated immunity, particularly involving class I human leukocyte antigen-restricted cytotoxic T lymphocytes (CTL), is believed to be crucial in combatting established HBV infection.

Class I human leukocyte antigen (HLA) molecules are expressed on the surface of almost all nucleated cells. CTL recognize peptide fragments, derived from intracellular processing of various antigens, in the form of a complex with class I HLA molecules. This recognition event then results in the destruction of the cell bearing the HLA-peptide complex directly or the activation of non-destructive mechanisms e.g., the production of interferon, that inhibit viral replication.

Several studies have emphasized the association between self-limiting acute hepatitis and multispecific CTL responses (Penna, A. *et al.*, *J. Exp. Med.* 174:1565, 1991; Nayersina, R. *et al.*, *J. Immunol.* 150:4659, 1993). Spontaneous and interferon-related clearance of chronic HBV infection is also associated with the resurgence of a vigorous CTL response (Guidotti, L. G. *et al.*, *Proc. Natl. Acad. Sci. USA* 91:3764, 1994). In all such cases the CTL responses are polyclonal, and specific for multiple viral proteins including the HBV envelope, core and polymerase antigens. By contrast, in patients with chronic hepatitis, the CTL activity is usually absent or weak, and antigenically restricted.

The crucial role of CTL in resolution of HBV infection has been further underscored by studies using HBV transgenic mice. Adoptive transfer of HBV-specific CTL into mice transgenic for the HBV genome resulted in suppression of virus replication. This effect was primarily mediated by a non-lytic, lymphokine-based mechanism (Guidotti, L. G. *et al.*, *Proc. Natl. Acad. Sci. USA* 91:3764, 1994; Guidotti, L.

G., Guilhot, S., and Chisari, F. V. *J. Virol.* 68:1265, 1994; Guidotti, L. G. *et al.*, *J. Virol.* 69:6158, 1995; Gilles, P. N., Fey, G., and Chisari, F. V., *J. Virol.* 66:3955, 1992).

As is the case for HLA class I restricted responses, HLA class II restricted T cell responses are usually detected in patients with acute hepatitis, and are absent or weak in patients with chronic infection (Chisari, F. V. and Ferrari, C., *Annu. Rev. Immunol.* 13:29, 1995). HLA Class II responses are tied to activation of helper T cells (HTLs) Helper T lymphocytes, which recognize Class II HLA molecules, may directly contribute to the clearance of HBV infection through the secretion of cytokines which suppress viral replication (Franco, A. *et al.*, *J. Immunol.* 159:2001, 1997). However, their primary role in disease resolution is believed to be mediated by inducing activation and expansion of virus-specific CTL and B cells.

In view of the heterogeneous immune response observed with HBV infection, induction of a multi-specific cellular immune response directed simultaneously against multiple epitopes appears to be important for the development of an efficacious vaccine against HBV. There is a need to establish vaccine embodiments that elicit immune responses that correspond to responses seen in patients that clear HBV infection. Epitope-based vaccines appear useful.

Upon development of appropriate technology, the use of epitope-based vaccines has several advantages over current vaccines. The epitopes for inclusion in such a vaccine are to be selected from conserved regions of viral or tumor-associated antigens, in order to reduce the likelihood of escape mutants. The advantage of an epitope-based approach over the use of whole antigens is that there is evidence that the immune response to whole antigens is directed largely toward variable regions of the antigen, allowing for immune escape due to mutations. Furthermore, immunosuppressive epitopes that may be present in whole antigens can be avoided with the use of epitope-based vaccines.

Additionally, with an epitope-based vaccine approach, there is an ability to combine selected epitopes (CTL and HTL) and additionally to modify the composition of the epitopes, achieving, for example, enhanced immunogenicity. Accordingly, the immune response can be modulated, as appropriate, for the target disease. Similar engineering of the response is not possible with traditional approaches.

Another major benefit of epitope-based immune-stimulating vaccines is their safety. The possible pathological side effects caused by infectious agents or whole protein antigens, which might have their own intrinsic biological activity, is eliminated.

An epitope-based vaccine also provides the ability to direct and focus an immune response to multiple selected antigens from the same pathogen. Thus, patient-by-patient variability in the immune response to a particular pathogen may be alleviated by inclusion of epitopes from multiple antigens from that pathogen in a vaccine composition. A  
 5 “pathogen” may be an infectious agent or a tumor associated molecule.

However, one of the most formidable obstacles to the development of broadly efficacious epitope-based immunotherapeutics has been the extreme polymorphism of HLA molecules. To date, effective non-genetically biased coverage of a population has been a task of considerable complexity; such coverage has required that epitopes be used  
 10 specific for HLA molecules corresponding to each individual HLA allele, therefore, impractically large numbers of epitopes would have to be used in order to cover ethnically diverse populations. There has existed a need to develop peptide epitopes that are bound by multiple HLA antigen molecules for use in epitope-based vaccines. The greater the number of HLA antigen molecules bound, the greater the breadth of  
 15 population coverage by the vaccine.

Furthermore, as described herein in greater detail, a need has existed to modulate peptide binding properties, for example so that peptides that are able to bind to multiple HLA antigens do so with an affinity that will stimulate an immune response. Identification of epitopes restricted by more than one HLA allele at an affinity that  
 20 correlates with immunogenicity is important to provide thorough population coverage, and to allow the elicitation of responses of sufficient vigor whereby the natural immune responses noted in self-limiting acute hepatitis, or of spontaneous clearance of chronic HBV infection is induced in a diverse segment of the population. Such a response can also target a broad array of epitopes. The technology disclosed herein provides for such  
 25 favored immune responses.

The information provided in this section is intended to disclose the presently understood state of the art as of the filing date of the present application. Information is included in this section which was generated subsequent to the priority date of this application. Accordingly, background in this section is not intended, in any way, to  
 30 delineate the priority date for the invention.

## II. SUMMARY OF THE INVENTION

This invention applies our knowledge of the mechanisms by which antigen is recognized by T cells, for example, to develop epitope-based vaccines directed towards

HBV. More specifically, this application communicates our discovery of specific epitope pharmaceutical compositions and methods of use in the prevention and treatment of HBV infection.

Upon development of appropriate technology, the use of epitope-based vaccines has several advantages over current vaccines, particularly when compared to the use of whole antigens in vaccine compositions. There is evidence that the immune response to whole antigens is directed largely toward variable regions of the antigen, allowing for immune escape due to mutations. The epitopes for inclusion in an epitope-based vaccine are selected from conserved regions of viral or tumor-associated antigens, which thereby reduces the likelihood of escape mutants. Furthermore, immunosuppressive epitopes that may be present in whole antigens can be avoided with the use of epitope-based vaccines.

An additional advantage of an epitope-based vaccine approach is the ability to combine selected epitopes (CTL and HTL), and further, to modify the composition of the epitopes, achieving, for example, enhanced immunogenicity. Accordingly, the immune response can be modulated, as appropriate, for the target disease. Similar engineering of the response is not possible with traditional approaches.

Another major benefit of epitope-based immune-stimulating vaccines is their safety. The possible pathological side effects caused by infectious agents or whole protein antigens, which might have their own intrinsic biological activity, is eliminated.

An epitope-based vaccine also provides the ability to direct and focus an immune response to multiple selected antigens from the same pathogen. Thus, patient-by-patient variability in the immune response to a particular pathogen may be alleviated by inclusion of epitopes from multiple antigens from that pathogen in a vaccine composition. A “pathogen” may be an infectious agent or a tumor associated molecule.

One of the most formidable obstacles to the development of broadly efficacious epitope-based immunotherapeutics, however, has been the extreme polymorphism of HLA molecules. To date, effective non-genetically biased coverage of a population has been a task of considerable complexity; such coverage has required that epitopes be used specific for HLA molecules corresponding to each individual HLA allele, therefore, impractically large numbers of epitopes would have to be used in order to cover ethnically diverse populations. Thus, there has existed a need to develop peptide epitopes that are bound by multiple HLA antigen molecules for use in epitope-based vaccines. The greater the number of HLA antigen molecules bound, the greater the breadth of population coverage by the vaccine.

Furthermore, as described herein in greater detail, a need has existed to modulate peptide binding properties, for example, so that peptides that are able to bind to multiple HLA antigens do so with an affinity that will stimulate an immune response.

Identification of epitopes restricted by more than one HLA allele at an affinity that correlates with immunogenicity is important to provide thorough population coverage, and to allow the elicitation of responses of sufficient vigor to prevent or clear an infection in a diverse segment of the population. Such a response can also target a broad array of epitopes. The technology disclosed herein provides for such favored immune responses.

In a preferred embodiment, epitopes for inclusion in vaccine compositions of the invention are selected by a process whereby protein sequences of known antigens are evaluated for the presence of motif or supermotif-bearing epitopes. Peptides corresponding to a motif- or supermotif-bearing epitope are then synthesized and tested for the ability to bind to the HLA molecule that recognizes the selected motif. Those peptides that bind at an intermediate or high affinity *i.e.*, an  $IC_{50}$  (or a  $K_D$  value) of 500 nM or less for HLA class I molecules or 1000 nM or less for HLA class II molecules, are further evaluated for their ability to induce a CTL or HTL response. Immunogenic peptides are selected for inclusion in vaccine compositions.

Supermotif-bearing peptides may additionally be tested for the ability to bind to multiple alleles within the HLA supertype family. Moreover, peptide epitopes may be analogued to modify binding affinity and/or the ability to bind to multiple alleles within an HLA supertype.

The invention also includes an embodiment comprising a method for monitoring immunogenic activity of a vaccine for HBV in a patient having a known HLA-type, the method comprising incubating a T lymphocyte sample from the patient with a peptide composition comprising an HBV epitope consisting essentially of an amino acid sequence described in Tables VI to Table XX or Table XXII which binds the product of at least one HLA allele present in said patient, and detecting for the presence of a T lymphocyte that binds to the peptide. In a preferred embodiment, the peptide comprises a tetrameric complex.

An alternative modality for defining the peptides in accordance with the invention is to recite the physical properties, such as length; primary, potentially secondary and/or tertiary structure; or charge, which are correlated with binding to a particular allele-specific HLA molecule or group of allele-specific HLA molecules. A further modality for defining peptides is to recite the physical properties of an HLA binding pocket, or

properties shared by several allele-specific HLA binding pockets (*e.g.* pocket configuration and charge distribution) and reciting that the peptide fits and binds to said pocket or pockets.

As will be apparent from the discussion below, other methods and embodiments are also contemplated. Further, novel synthetic peptides produced by any of the methods described herein are also part of the invention.

### III. BRIEF DESCRIPTION OF THE FIGURES

Figure 1: Figure 1 provides a graph of total frequency of genotypes as a function of the number of HBV candidate epitopes bound by HLA-A and B molecules, in an average population.

Figure 2: Figure 2 Illustrates the Position of Peptide Epitopes in Experimental Model Minigene Constructs

### IV. DETAILED DESCRIPTION OF THE INVENTION

The peptides and corresponding nucleic acid compositions of the present invention are useful for stimulating an immune response to HBV either by stimulating the production of CTL or HTL responses. The peptides, which are derived directly or indirectly from native HBV amino acid sequences, are able to bind to HLA molecules and stimulate an immune response to HBV. The complete polyprotein sequence from HBV and its variants can be obtained from Genbank. Peptides can also be readily determined from sequence information that may subsequently be discovered for heretofore unknown variants of HBV as will be clear from the disclosure provided below.

The peptides of the invention have been identified in a number of ways, as will be discussed below. Further, analog peptides have been derived and the binding activity for HLA molecules modulated by modifying specific amino acid residues to create peptide analogs exhibiting altered immunogenicity. Further, the present invention provides compositions and combinations of compositions that enable epitope-based vaccines that are capable of interacting with multiple HLA antigens to provide broader population coverage than prior vaccines.

#### IV.A. Definitions

The invention can be better understood with reference to the following definitions, which are listed alphabetically.



“Cross-reactive binding” indicates that a peptide is bound by more than one HLA molecule; a synonym is degenerate binding.

A "cryptic epitope" elicits a response by immunization with an isolated peptide, but the response is not cross-reactive *in vitro* when intact whole protein which comprises the epitope is used as an antigen.

A "dominant epitope" is an epitope that induces an immune response upon immunization with a whole native antigen.. (See, *e.g.*, Sercarz, *et al.*, *Annu. Rev. Immunol.* 11:729766 (1993)) Such a response is cross-reactive *in vitro* with an isolated peptide epitope.

With regard to a particular amino acid sequence, an “epitope” is a set of amino acid residues which is involved in recognition by a particular immunoglobulin, or in the context of T cells, those residues necessary for recognition by T cell receptor proteins and/or Major Histocompatibility Complex (MHC) receptors. In an immune system setting, *in vivo* or *in vitro*, an epitope is the collective features of a molecule, such as primary, secondary and tertiary peptide structure, and charge, that together form a site recognized by an immunoglobulin, T cell receptor or HLA molecule.

“Human Leukocyte Antigen” or “HLA” is a human class I or class II Major Histocompatibility Complex (MHC) protein (*see*, Stites, *et al.*, IMMUNOLOGY, 8<sup>TH</sup> ED., Lange Publishing, Los Altos, CA (1994).

An "HLA supertype or family", as used herein, describes sets of HLA molecules grouped on the basis of shared peptide-binding specificities. HLA class I molecules that share somewhat similar binding affinity for peptides bearing certain amino acid motifs are grouped into HLA superotypes. The terms HLA superfamily, HLA supertype family, and HLA xx-like supertype molecules (where xx denotes a particular HLA type) are synonyms.

Throughout this disclosure, results are expressed in terms of “IC<sub>50</sub>'s.” IC<sub>50</sub> is the concentration of peptide in a binding assay at which 50% inhibition of binding of a reference peptide is observed. Given the conditions in which the assays are run (*i.e.*, limiting HLA proteins and labeled peptide concentrations), these values approximate K<sub>D</sub> values. It should be noted that IC<sub>50</sub> values can change, often dramatically, if the assay conditions are varied, and depending on the particular reagents used (*e.g.*, HLA preparation, *etc.*). For example, excessive concentrations of HLA molecules will increase the apparent measured IC<sub>50</sub> of a given ligand.

Assays for determining binding are described in detail in PCT publications WO 94/20127 and WO 94/03205. Alternatively, binding is expressed relative to a reference peptide. As a particular assay becomes more, or less, sensitive, the  $IC_{50}$ 's of the peptides tested may change somewhat. However, the binding relative to the reference peptide will not significantly change. For example, in an assay run under conditions such that the  $IC_{50}$  of the reference peptide increases 10-fold, the  $IC_{50}$  values of the test peptides will also shift approximately 10-fold. Therefore, to avoid ambiguities, the assessment of whether a peptide is a good, intermediate, weak, or negative binder is generally based on its  $IC_{50}$ , relative to the  $IC_{50}$  of a standard peptide.

Binding may also be determined using other assays including, for example, inhibition of antigen presentation (Sette *et al.*, *J. Immunol.* 141:3893, 1991), *in vitro* assembly assays (Townsend *et al.*, *Cell* 62:285, 1990), measures of dissociations rates (Parker *et al.*, *J. Immunol.* 149:1896-1904, 1992), and FACS-based assays using mutated cells, such as RMA.S (Melief, *et al.*, *Eur. J. Immunol.* 21:2963, 1991).

As used herein, high affinity with respect to HLA class I molecules is defined as binding with an  $IC_{50}$  or  $K_D$  value of less than 50 nM; intermediate affinity is binding with an  $IC_{50}$  (or  $K_D$ ) of between about 50 and about 500 nM. High affinity with respect to binding to HLA class II molecules is defined as binding with an  $IC_{50}$  or  $K_D$  value of less than 100 nM; intermediate affinity is binding with an  $IC_{50}$  or  $K_D$  of between about 100 and about 1000 nM.

The terms "identical" or percent "identity," in the context of two or more peptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using a sequence comparison algorithms or by manual alignment and visual inspection.

An "immunogenic peptide" or "peptide epitope" is a peptide which comprises an allele-specific motif or supermotif such that the peptide will bind an HLA molecule and induce a CTL and/or HTL response. Thus, immunogenic peptides of the invention are capable of binding to an appropriate HLA molecule and thereafter inducing a cytotoxic T cell response, or a helper T cell response, to the antigen from which the immunogenic peptide is derived.

The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany the material as it is found in its native state. Thus, isolated peptides in accordance with the invention

preferably do not contain materials normally associated with the peptides in their *in situ* environment.

“Major Histocompatibility Complex” or “MHC” is a cluster of genes that plays a role in control of the cellular interactions responsible for physiologic immune responses.

5 In humans, the MHC complex is also known as the HLA complex. For a detailed description of the MHC and HLA complexes, see, Paul, FUNDAMENTAL IMMUNOLOGY, 3<sup>RD</sup> ED., Raven Press, New York, 1993.

The term "motif" refers to the pattern of residues in a peptide of defined length, usually a peptide of from about 8 to about 13 amino acids for a class I HLA motif and  
10 from about 6 to about 25 amino acids for a class II HLA motif, which is recognized by a particular HLA molecule. Peptide motifs are typically different for each protein encoded by each human HLA allele and differ in the pattern of the primary and secondary anchor residues.

A "negative binding residue" or “deleterious residue” is an amino acid which, if  
15 present at certain positions (typically not primary anchor positions) of a peptide epitope, results in decreased binding affinity of the peptide for the peptide’s corresponding HLA molecule. Any residue that is not “deleterious” is a “non-deleterious” residue.

The term "peptide" is used interchangeably with "oligopeptide" in the present specification to designate a series of residues, typically L-amino acids, connected one to  
20 the other, typically by peptide bonds between the  $\alpha$ -amino and carboxyl groups of adjacent amino acids. The preferred CTL-inducing oligopeptides of the invention are 13 residues or less in length and usually consist of between about 8 and about 11 residues, preferably 9 or 10 residues. The preferred HTL-inducing oligopeptides are less than about 50 residues in length and usually consist of between about 6 and about 30 residues,  
25 more usually between about 12 and 25, and often between about 15 and 20 residues.

“Pharmaceutically acceptable” refers to a non-toxic, inert, and physiologically compatible composition.

A "primary anchor residue" is an amino acid at a specific position along a peptide sequence which is understood to provide a contact point between the immunogenic  
30 peptide and the HLA molecule. One to three, usually two, primary anchor residues within a peptide of defined length generally defines a “motif” for an immunogenic peptide. These residues are understood to fit in close contact with peptide binding grooves of an HLA molecule, with their side chains buried in specific pockets of the binding grooves themselves. In one embodiment, the primary anchor residues are located

at position 2 (from the amino terminal position) and at the carboxyl terminal position of a 9 residue peptide in accordance with the invention. The primary anchor positions for each motif and supermotif are set forth in Table I. For example, analog peptides can be created by altering the presence or absence of particular residues in these primary anchor positions. Such analogs are used to finely modulate the binding affinity of a peptide comprising a particular motif or supermotif.

“Promiscuous recognition” is where a distinct peptide is recognized by the same T cell clone in the context of multiple HLA molecules. Promiscuous binding is synonymous with cross-reactive binding.

A “protective immune response” or “therapeutic immune response” refers to a CTL and/or an HTL response to an antigen from an infectious agent or a tumor antigen from which an immunogenic peptide is derived, and thereby preventing or at least partially arresting disease symptoms or progression. The immune response may also include an antibody response which has been facilitated by the stimulation of helper T cells.

The term “residue” refers to an amino acid or amino acid mimetic incorporated into an oligopeptide by an amide bond or amide bond mimetic.

A “secondary anchor residue” is an amino acid at a position other than a primary anchor position in a peptide which may influence peptide binding. A secondary anchor residue occurs at a significantly higher frequency amongst bound peptides than would be expected by random distribution of amino acids at one position. The secondary anchor residues are said to occur at “secondary anchor positions.” A secondary anchor residue can be identified as a residue which is present at a higher frequency among high affinity binding peptides, or a residue otherwise associated with high affinity binding. For example, analog peptides can be created by altering the presence or absence of particular residues in these secondary anchor positions. Such analogs are used to finely modulate the binding affinity of a peptide comprising a particular motif or supermotif.

A “subdominant epitope” is an epitope which evokes little or no response upon immunization with whole antigens which comprise the epitope, but for which a response can be obtained by immunization with an isolated peptide, and this response (unlike the case of cryptic epitopes) is detected when whole protein is used to recall the response *in vitro* or *in vivo*.

A "supermotif" is a peptide binding specificity shared by HLA molecules encoded by two or more HLA alleles. A supermotif-bearing epitope is preferably recognized with high or intermediate affinity (as defined herein) by two or more HLA antigens.

5 "Synthetic peptide" refers to a peptide that is not naturally occurring, but is man-made using such methods as chemical synthesis or recombinant DNA technology.

The nomenclature used to describe peptide compounds follows the conventional practice wherein the amino group is presented to the left (the N-terminus) and the carboxyl group to the right (the C-terminus) of each amino acid residue. When amino acid residue positions are referred to in a peptide epitope they are numbered in an amino  
10 to carboxyl direction with position one being the position closest to the amino terminal. In the formulae representing selected specific embodiments of the present invention, the amino- and carboxyl-terminal groups, although not specifically shown, are in the form they would assume at physiologic pH values, unless otherwise specified. In the amino acid structure formulae, each residue is generally represented by standard three letter or  
15 single letter designations. The L-form of an amino acid residue is represented by a capital single letter or a capital first letter of a three-letter symbol, and the D-form for those amino acids having D-forms is represented by a lower case single letter or a lower case three letter symbol. Glycine has no asymmetric carbon atom and is simply referred to as "Gly" or G. Symbols for the amino acids are shown below.

Single Letter Symbol	Three Letter Symbol	Amino Acids
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic Acid
E	Glu	Glutamic Acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

#### IV.B. Stimulation of CTL and HTL responses against HBV

The mechanism by which T cells recognize antigens has been delineated during the past ten years. Based on our new understanding of the immune system we have generated efficacious peptide epitope vaccine compositions that can induce a therapeutic or prophylactic immune response to HBV infection in a broad population. For an understanding of the value and efficacy of the claimed compositions, a brief review of the technology is provided.

A complex of an HLA molecule and a peptidic antigen acts as the ligand recognized by HLA-restricted T cells (Buus, S. *et al.*, *Cell* 47:1071, 1986; Babbitt, B. P. *et al.*, *Nature* 317:359, 1985; Townsend, A., and Bodmer, H., *Annu. Rev. Immunol.* 7:601,

1989; Germain, R. N., *Annu. Rev. Immunol.* 11:403, 1993). Through the study of single amino acid substituted antigen analogs and the sequencing of endogenously bound, naturally processed peptides, critical residues that correspond to motifs required for specific binding to HLA antigen molecules have been identified and are described here and set forth in Tables I, II, and III (see also, e.g., Sette, A. and Grey, H. M., *Curr. Opin. Immunol.* 4:79, 1992; Sinigaglia, F. and Hammer, J., *Curr. Biol.* 6:52, 1994; Engelhard, V. H., *Curr. Opin. Immunol.* 6:13, 1994). Furthermore, x-ray crystallographic analysis of HLA-peptide complexes has revealed pockets within the peptide binding cleft of HLA molecules which accommodate allele-specific residues borne by peptide ligands; these residues in turn determine the HLA binding capacity of the peptides in which they are present (Brown, J. H. *et al.*, *Nature* 364:33, 1993; Guo, H. C. *et al.*, *Proc. Natl. Acad. Sci. USA* 90:8053, 1993; Guo, H. C. *et al.*, *Nature* 360:364, 1992; Silver, M. L. *et al.*, *Nature* 360:367, 1992; Matsumura, M. *et al.*, *Science* 257:927, 1992; Madden *et al.*, *Cell* 70:1035, 1992; Fremont, D. H. *et al.*, *Science* 257:919, 1992; Saper, M. A., Bjorkman, P. J. and Wiley, D. C., *J. Mol. Biol.* 219:277, 1991).

Accordingly, the definition of class I and class II allele-specific HLA binding motifs or class I supermotifs allows identification of regions within a protein that have the potential of binding particular HLA antigens (see also e.g., Sette, A. and Grey, H. M., *Curr. Opin. Immunol.* 4:79, 1992; Sinigaglia, F. and Hammer, J., *Curr. Biol.* 6:52, 1994; Engelhard, V. H., *Curr. Opin. Immunol.* 6:13, 1994; Kast, W. M. *et al.*, *J. Immunol.*, 152:3904, 1994).

Furthermore, a variety of assays to quantify the affinity of interaction between peptide and HLA have also been established. Such assays include, for example, measures of IC<sub>50</sub> values, inhibition of antigen presentation (Sette *et al.*, *J. Immunol.* 141:3893, 1991), *in vitro* assembly assays (Townsend *et al.*, *Cell* 62:285, 1990), measures of dissociations rates (Parker *et al.*, *J. Immunol.* 149:1896-1904, 1992), and FACS-based assays using mutated cells, such as RMA.S (Melief, *et al.*, *Eur. J. Immunol.* 21:2963, 1991).

The present inventors have found that the correlation of binding affinity with immunogenicity is an important factor to be considered when evaluating candidate peptides. Thus, by a combination of motif searches and HLA-peptide binding assays, candidates for epitope-based vaccines have been identified. After determining their binding affinity, additional confirmatory work can be performed to select, amongst these vaccine candidates, epitopes with preferred characteristics in terms of antigenicity and

immunogenicity. Various strategies can be utilized to evaluate immunogenicity, including:

1) Evaluation of primary T cell cultures from normal individuals (Wentworth, P. A. *et al.*, *Mol. Immunol.* 32:603, 1995; Celis, E. *et al.*, *Proc. Natl. Acad. Sci. USA* 91:2105, 1994; Tsai, V. *et al.*, *J. Immunol.* 158:1796, 1997; Kawashima, I. *et al.*, *Human Immunol.* 59:1, 1998); This procedure involves the stimulation of PBL from normal subjects with a test peptide in the presence of antigen presenting cells *in vitro* over a period of several weeks. T cells specific for the peptide become activated during this time and are detected using a  $^{51}\text{Cr}$ -release assay involving peptide sensitized target cells.

2) Immunization of HLA transgenic mice (Wentworth, P. A. *et al.*, *J. Immunol.* 26:97, 1996; Wentworth, P. A. *et al.*, *Int. Immunol.* 8:651, 1996; Alexander, J. *et al.*, *J. Immunol.* 159:4753, 1997); In this method, peptides in incomplete Freund's adjuvant are administered subcutaneously to HLA transgenic mice. Several weeks following immunization, splenocytes are removed and cultured *in vitro* in the presence of test peptide for approximately one week. Peptide-specific T cells are detected using a  $^{51}\text{Cr}$ -release assay involving peptide sensitized target cells and target cells expressing endogenously generated antigen.

3) Demonstration of recall T cell responses from immune individuals who have recovered from infection, and/or from chronically infected patients (Rehermann, B. *et al.*, *J. Exp. Med.* 181:1047, 1995; Doolan, D. L. *et al.*, *Immunity* 7:97, 1997; Bertoni, R. *et al.*, *J. Clin. Invest.* 100:503, 1997; Threlkeld, S. C. *et al.*, *J. Immunol.* 159:1648, 1997; Diepolder, H. M. *et al.*, *J. Virol.* 71:6011, 1997). In applying this strategy, recall responses were detected by culturing PBL from subjects that had been naturally exposed to the antigen, for instance through infection, and thus had generated an immune response "naturally". PBL from subjects were cultured *in vitro* for 1-2 weeks in the presence of test peptide plus antigen presenting cells (APC) to allow activation of "memory" T cells, as compared to "naive" T cells. At the end of the culture period, T cell activity is detected using assays for T cell activity including  $^{51}\text{Cr}$  release involving peptide-sensitized targets, T cell proliferation or lymphokine release.

The following describes the peptide epitopes and corresponding nucleic acids of the invention.



#### IV.C. Binding Affinity of Peptide Epitopes for HLA Molecules

As indicated herein, the large degree of HLA polymorphism is an important factor to be taken into account with the epitope-based approach to vaccine development. To address this factor, epitope selection encompassing identification of peptides capable of binding at high or intermediate affinity to multiple HLA molecules is preferably utilized, most preferably these epitopes bind at high or intermediate affinity to two or more allele specific HLA molecules.

CTL-inducing peptides of interest for vaccine compositions preferably include those that have an  $IC_{50}$  or binding affinity value for class I HLA molecules of 500 nM or less. HTL-inducing peptides preferably include those that have an  $IC_{50}$  or binding affinity value for class II HLA molecules of 1000 nM or less. For example, peptide binding is assessed by testing the capacity of a candidate peptide to bind to a purified HLA molecule *in vitro*. Peptides exhibiting high or intermediate affinity are then considered for further analysis. Selected peptides are tested on other members of the supertype family. In preferred embodiments, peptides that exhibit cross-reactive binding are then used in vaccines or in cellular screening analyses.

As disclosed herein, high HLA binding affinity is correlated with greater immunogenicity. Greater immunogenicity can be manifested in several different ways. Immunogenicity corresponds to whether an immune response is elicited at all, and to the vigor of any particular response. For example, a peptide might elicit an immune response in a diverse array of the population, yet in no instance produce a vigorous response. In accordance with these principles, close to 90% of high binding peptides have been found to be immunogenic, as contrasted with about 50% of the peptides which bind with intermediate affinity. Moreover, higher binding affinity peptides leads to more vigorous immunogenic responses. As a result, less peptide is required to elicit a similar biological effect if a high affinity binding peptide is used. Thus, in preferred embodiments of the invention, high binding epitopes are particularly desired.

The relationship between binding affinity for HLA class I molecules and immunogenicity of discrete peptide epitopes on bound antigens has been determined for the first time in the art by the present inventors. The correlation between binding affinity and immunogenicity was analyzed in two different experimental approaches (Sette, *et al.*, *J. Immunol.* 153:5586-5592, 1994). In the first approach, the immunogenicity of potential epitopes ranging in HLA binding affinity over a 10,000-fold range was analyzed in HLA-A\*0201 transgenic mice. In the second approach, the antigenicity of

approximately 100 different hepatitis B virus (HBV)-derived potential epitopes, all carrying A\*0201 binding motifs, was assessed by using PBL (peripheral blood lymphocytes) of acute hepatitis patients. Pursuant to these approaches, it was determined that an affinity threshold of approximately 500 nM (preferably an IC<sub>50</sub> value of 500 nM or less) determines the capacity of a peptide epitope to elicit a CTL response. These data are true for class I binding affinity measurements for naturally processed peptides and for synthesized T cell epitopes. These data also indicate the important role of determinant selection in the shaping of T cell responses.

An affinity threshold associated with immunogenicity in the context of HLA class II DR molecules has also been delineated (Southwood *et al. J. Immunology* 160:3363-3373, 1998, and U.S.S.N 60/087192 filed 5/29/98). In order to define a biologically significant threshold of DR binding affinity, a database of the binding affinities of 32 DR-restricted epitopes for their restricting element was compiled. In approximately half of the cases (15 of 32 epitopes), DR restriction was associated with high binding affinities, *i.e.*, binding affinities of with an IC<sub>50</sub> value of 100 nM or less. In the other half of the cases (16 of 32), DR restriction was associated with intermediate affinity (binding affinities in the 100-1000 nM range). In only one of 32 cases was DR restriction associated with an IC<sub>50</sub> of 1000 nM or greater. Thus, 1000 nM can be defined as an affinity threshold associated with immunogenicity in the context of DR molecules.

The binding affinity of peptides for HLA molecules can be determined as described in Example 1, below.

#### IV.D. Peptide Epitope Binding Motifs and Supermotifs

In the past few years evidence has accumulated to demonstrate that a large fraction of HLA class I, and possibly class II molecules can be classified into a relatively few supertypes characterized by largely overlapping peptide binding repertoires, and consensus structures of the main peptide binding pockets.

For HLA molecule pocket analyses, the residues comprising the B and F pockets of HLA class I molecules as described in crystallographic studies (Guo, H. C. *et al.*, *Nature* 360:364, 1992; Saper, M. A., Bjorkman, P. J. and Wiley, D. C., *J. Mol. Biol.* 219:277, 1991; Madden, D. R., Garboczi, D. N. and Wiley, D. C., *Cell* 75:693, 1993), have been compiled from the database of Parham, *et al.* (Parham, P., Adams, E. J., and Arnett, K. L., *Immunol. Rev.* 143:141, 1995). In these analyses, residues 9, 45, 63, 66, 67, 70, and 99 were considered to make up the B pocket, and to determine the specificity

for the residue in the second position of peptide ligands. Similarly, residues 77, 80, 81, and 116 were considered to determine the specificity of the F pocket, and to determine the specificity for the C-terminal residue of a peptide ligand bound by the HLA molecule.

Through the study of single amino acid substituted antigen analogs and the sequencing of endogenously bound, naturally processed peptides, critical residues required for allele-specific binding to HLA molecules have been identified. The presence of these residues correlates with binding affinity for HLA molecules. The identification of motifs and/or supermotifs that correlate with high and intermediate affinity binding is an important issue with respect to the identification of immunogenic peptide epitopes for the inclusion in a vaccine. Kast *et al.* (*J. Immunol.* 152:3904-3912, 1994) have shown that motif-bearing peptides account for 90% of the epitopes that bind to allele-specific HLA class I molecules. In this study all possible peptides of 9 amino acids in length and overlapping by eight amino acids (240 peptides), which cover the entire sequence of the E6 and E7 proteins of human papillomavirus type 16, were evaluated for binding to five allele-specific HLA molecules that are expressed at high frequency among different ethnic groups. This unbiased set of peptides allowed an evaluation of the predictive value of HLA class I motifs. From the set of 240 peptides, 22 peptides were identified that bound to an allele-specific HLA molecules with high or intermediate affinity. Of these 22 peptides, 20, (*i.e.* 91%), were motif-bearing. Thus, this study demonstrates the value of motifs for the identification of peptide epitopes for inclusion in a vaccine: application of motif-based identification techniques eliminates screening of 90% of the potential epitopes.

Such peptide epitopes are identified in the Tables described below. The Tables for the HLA class I epitopes include over 90% of the peptides that will bind to an allele-specific HLA class I molecule with intermediate or high affinity.

Peptides of the present invention may also include epitopes that bind to MHC class II DR molecules. A significant difference between class I and class II HLA molecules is that, although a stringent size restriction exists for peptide binding to class I molecules, a greater degree of heterogeneity in both sizes and binding frame positions of the motif, relative to the N and C termini of the peptide, can be demonstrated for class II peptide ligands. This increased heterogeneity is due to the structure of the class II-binding groove which, unlike its class I counterpart, is open at both ends.

Crystallographic analysis of DRB\*0101-peptide complexes (*see, e.g.*, Madden, D.R. *Ann. Rev. Immunol.* 13:587, 1995) showed that the residues occupying position 1 and position

6 of peptides complexed with DRB\*0101 engage two complementary pockets on the DRBa\*0101 molecules, with the P1 position corresponding to the most crucial anchor residue and the deepest hydrophobic pocket. Other studies have also pointed to the P6 position as a crucial anchor residue for binding to various other DR molecules.

5 Thus, peptides of the present invention are identified by any one of several HLA-specific amino acid motifs(*see, e.g.*, Tables I-III). If the presence of the motif corresponds to the ability to bind several allele-specific HLA antigens it is referred to as a supermotif. The allele-specific HLA molecules that bind to peptides that possess a particular amino acid supermotif are collectively referred to as an HLA “supertype.”

10 The peptide motifs and supermotifs described below provide guidance for the identification and use of peptides in accordance with the invention.

Examples of peptide epitopes bearing the respective supermotif or motif are included in Tables as designated in the description of each motif or supermotif. The Tables include a binding affinity ratio listing for some of the peptide epitopes. The ratio  
15 may be converted to IC<sub>50</sub> by using the following formula: IC<sub>50</sub> of the standard peptide/ratio = IC<sub>50</sub> of the test peptide (i.e. the peptide epitope). The IC<sub>50</sub> values of standard peptides used to determine binding affinities for Class I peptides are shown in Table IV. The IC<sub>50</sub> values of standard peptides used to determine binding affinities for Class II peptides are shown in Table V. The peptides used as standards for the binding  
20 assay are examples of standards; alternative standard peptides can also be used when performing such an analysis.

To obtain the peptide epitope sequences listed in each Table, protein sequence data from twenty HBV strains (HPBADR, HPBADR1CG, HPBADRA, HPBADRC, HPBADRCG, HPBCGADR, HPBVADRM, HPBADW, HPBADW1, HPBADW2,  
25 HPBADW3, HPBADWZ, HPBHEPB, HPBVADW2, HPBAYR, HPBV, HPBVAYWC, HPBVAYWCI, NAD HPBVAYWE) were evaluated for the presence of the designated supermotif or motif. Peptide epitopes were also selected on the basis of their conservancy. A criterion for conservancy requires that the entire sequence of a peptide be totally conserved in 75% of the sequences available for a specific protein. The percent  
30 conservancy of the selected peptide epitopes is indicated on the Tables. The frequency, *i.e.* the number of strains of the 20 strains in which the peptide sequence was identified, is also shown. The “1<sup>st</sup> position” column in the Tables designates the amino acid position of the HBV protein that corresponds to the first amino acid residue of the epitope. “Number of amino acids” indicates the number of residues in the epitope sequence.

### HLA Class I Motifs Indicative of CTL Inducing Peptide Epitopes:

The primary anchor residues of the HLA class I peptide epitope supermotifs and motifs delineated below are summarized in Table I. The HLA class I motifs set out in  
 5 Table I(a) are those most particularly relevant to the invention claimed here. Primary and secondary anchor positions are summarized in Table II. Allele-specific HLA molecules that comprise HLA class I supertype families are listed in Table VI.

#### IV.D1. HLA-A1 supermotif

10 The HLA-A1 supermotif is characterized by the presence in peptide ligands of a small (T or S) or hydrophobic (L, I, V, M, or F) primary anchor residue in position 2, and an aromatic (Y, F, or W) primary anchor residue at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind to the A1 supermotif (*i.e.*, the HLA-A1 supertype) is comprised of at least A\*0101, A\*2601, A\*2602, A\*2501, and  
 15 A\*3201 (*see, e.g.*, DiBrino, M. *et al.*, *J. Immunol.* 151:5930, 1993; DiBrino, M. *et al.*, *J. Immunol.* 152:620, 1994; Kondo, A. *et al.*, *Immunogenetics* 45:249, 1997.). Other allele-specific HLA molecules predicted to be members of the A1 superfamily are shown in Table VI. Peptides binding to each of the individual HLA proteins can be modulated by substitutions at primary anchor positions, preferably choosing respective residues  
 20 specified for the supermotif.

Representative peptide epitopes that comprise the A1 supermotif are set forth on the attached Table VII.

#### IV.D.2. HLA-A2 supermotif

25 Primary anchor specificities for allele-specific HLA A2.1 molecules (Falk *et al.*, *Nature* 351:290-296, 1991; Hunt *et al.*, *Science* 255:1261-1263, 1992) and cross-reactive binding within the HLA A2 family (Fruci *et al.*, *Human Immunol.* 38:187-192, 1993; Tanigaki *et al.*, *Human Immunol.* 39:155-162, 1994) have been described. The present inventors have defined additional primary anchor residues that determine cross-reactive  
 30 binding to multiple allele-specific HLA A2 molecules (Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995). The HLA-A2 supermotif comprises peptide ligands with L, I, V, M, A, T, or Q as primary anchor residues at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope.

The corresponding family of HLA molecules (*i.e.*, the HLA-A2 supertype that binds these peptides) is comprised of at least: A\*0201, A\*0202, A\*0203, A\*0204, A\*0205, A\*0206, A\*0207, A\*0209, A\*0214, A\*6802, and A\*6901. Other allele-specific HLA molecules predicted to be members of the A2 superfamily are shown in Table VI.

- 5 As explained in detail below, binding to each of the individual allele-specific HLA molecules can be modulated by substitutions at the primary anchor and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

- 10 Representative peptide epitopes that comprise an A2 supermotif are set forth on the attached Table VIII. The motifs comprising the primary anchor residues V, A, T, or Q at position 2 and L, I, V, A, or T at the C-terminal position are those most particularly relevant to the invention claimed herein.

#### IV.D.3. HLA-A3 supermotif

- 15 The HLA-A3 supermotif is characterized by the presence in peptide ligands of A, L, I, V, M, S, or, T as a primary anchor at position 2, and a positively charged residue, R or K, at the C-terminal position of the epitope (*e.g.*, in position 9 of 9-mers). Exemplary members of the corresponding family of HLA molecules (the HLA-A3 supertype) that bind the A3 supermotif include at least: A\*0301, A\*1101, A\*3101, A\*3301, and A\*6801. Other allele-specific HLA molecules predicted to be members of the A3
- 20 superfamily are shown in Table VI. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions of amino acids at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

- 25 Representative peptide epitopes that comprise the A3 supermotif are set forth on the attached Table IX.

#### IV.D.4. HLA-A24 supermotif

- 30 The HLA-A24 supermotif is characterized by the presence in peptide ligands of an aromatic (F, W, or Y) residue as a primary anchor in position 2, and a hydrophobic (Y, F, L, I, V, or M) residue as primary anchor at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind to the A24 supermotif (*i.e.*, the A24 supertype) includes at least A\*2402, A\*3001, and A\*2301. Other allele-specific HLA molecules predicted to be members of the A24 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by

substitutions at primary anchor positions, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that comprise the A24 supermotif are set forth on the attached Table X.

5

#### IV.D.5. HLA-B7 supermotif

The HLA-B7 supermotif is characterized by peptides bearing proline in position 2 as a primary anchor, and a hydrophobic or aliphatic amino acid (L, I, V, M, A, F, W, or Y) as the primary anchor at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind the B7 supermotif (*i.e.*, the HLA-B7 supertype) is comprised of at least twenty six HLA-B proteins including: B\*0702, B\*0703, B\*0704, B\*0705, B\*1508, B\*3501, B\*3502, B\*3503, B\*3504, B\*3505, B\*3506, B\*3507, B\*3508, B\*5101, B\*5102, B\*5103, B\*5104, B\*5105, B\*5301, B\*5401, B\*5501, B\*5502, B\*5601, B\*5602, B\*6701, and B\*7801 (*see, e.g.*, Sidney, *et al.*, *J. Immunol.* 154:247, 1995; Barber, *et al.*, *Curr. Biol.* 5:179, 1995; Hill, *et al.*, *Nature* 360:434, 1992; Rammensee, *et al.*, *Immunogenetics* 41:178, 1995). Other allele-specific HLA molecules predicted to be members of the B7 superfamily are shown in Table VI. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that contain the B7 supermotif are set forth on the attached Table XI.

#### IV.D.6. HLA-B27 supermotif

The HLA-B27 supermotif is characterized by the presence in peptide ligands of a positively charged (R, H, or K) residue as a primary anchor at position 2, and a hydrophobic (F, Y, L, W, M, I, A, or V) residue as a primary anchor at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA molecules that bind to the B27 supermotif (*i.e.*, the B27 supertype) include at least B\*1401, B\*1402, B\*1509, B\*2702, B\*2703, B\*2704, B\*2705, B\*2706, B\*3801, B\*3901, B\*3902, and B\*7301. Other allele-specific HLA molecules predicted to be members of the B27 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary anchor positions, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that comprise the B27 supermotif are set forth on the attached Table XII.

#### IV.D.7. HLA-B44 supermotif

5           The HLA-B44 supermotif is characterized by the presence in peptide ligands of negatively charged (D or E) residues as a primary anchor in position 2, and hydrophobic residues (F, W, Y, L, I, M, V, or A) as a primary anchor at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA molecules that bind to the B44 supermotif (*i.e.*, the B44 supertype) include at least: B\*1801, B\*1802, B\*3701,  
10   B\*4001, B\*4002, B\*4006, B\*4402, B\*4403, and B\*4006. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary anchor positions; preferably choosing respective residues specified for the supermotif.

#### IV.D.8. HLA-B58 supermotif

15           The HLA-B58 supermotif is characterized by the presence in peptide ligands of a small aliphatic residue (A, S, or T) as a primary anchor residue at position 2, and an aromatic or hydrophobic residue (F, W, Y, L, I, V, M, or A) as a primary anchor residue at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA molecules that bind to the B58 supermotif (*i.e.*, the B58 supertype)  
20   include at least: B\*1516, B\*1517, B\*5701, B\*5702, and B\*5801. Other allele-specific HLA molecules predicted to be members of the B58 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary anchor positions, preferably choosing respective residues specified for the supermotif.

25           Representative peptide epitopes that comprise the B58 supermotif are set forth on the attached Table XIII.

#### IV.D.9. HLA-B62 supermotif

30           The HLA-B62 supermotif is characterized by the presence in peptide ligands of the polar aliphatic residue Q or a hydrophobic aliphatic residue (L, V, M, or I) as a primary anchor in position 2, and a hydrophobic residue (F, W, Y, M, I, V, L, or A) as a primary anchor at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA molecules that bind to the B62 supermotif (*i.e.*, the B62 supertype) include at least: B\*1501, B\*1502, B\*1513, and B5201. Other allele-specific



HLA molecules predicted to be members of the B62 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary anchor positions, preferably choosing respective residues specified for the supermotif.

- 5           Representative peptide epitopes that comprise the B62 supermotif are set forth on the attached Table XIV.

#### IV.D.10. HLA-A1 motif

- 10           The allele-specific HLA-A1 motif is characterized by the presence in peptide ligands of T, S, or M as a primary anchor residue at position 2 and the presence of Y as a primary anchor residue at the C-terminal position of the epitope. An alternative allele-specific A1 motif (*i.e.*, a “submotif”) is characterized by a primary anchor residue at position 3 rather than position 2. This submotif is characterized by the presence of D, E, A, or S as a primary anchor residue in position 3, and a Y as a primary anchor residue at  
15           the C-terminal position of the epitope. Peptide binding to HLA A1 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

- 20           Representative peptide epitopes that comprise either A1 motif are set forth on the attached Table XV. Those epitopes comprising T, S, or M at position 2 and Y at the C-terminal position are also included in the listing of HLA-A1 supermotif-bearing peptide epitopes listed in Table VII.

#### IV.D.11. HLA-A2.1 motif

- 25           An allele-specific HLA-A2.1 motif was first determined to be characterized by the presence in peptide ligands of L or M as a primary anchor residue in position 2, and L or V as a primary anchor residue at the C-terminal position of a 9 amino acid epitope (Falk *et al.*, *Nature* 351:290-296, 1991). Furthermore, the A2.1 motif was determined to further comprise an I at position 2 and I or A at the C-terminal position of a nine amino acid peptide (Hunt *et al.*, *Science* 255:1261-1263, March 6, 1992). Additionally, the A2.1  
30           allele-specific motif has been found to comprise a T at the C-terminal position (Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994). Subsequently, the A2.1 allele-specific motif has been defined by the present inventors to additionally comprise V, A, T, or Q as a primary anchor residue at position 2, and M as a primary anchor residue at the C-terminal position of the epitope. Thus, the HLA-A2.1 motif comprises peptide ligands with L, I, V, M, A,

T, or Q as primary anchor residues at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope. The preferred and tolerated residues that characterize the primary anchor positions of the HLA-A2.1 motif are identical to the preferred residues of the A2 supermotif. (for reviews of relevant data, see, *e.g.*, Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Sidney *et al.*, *Immunol. Today* 17:261-266, 1996; Sette and Sidney, *Curr. Opin. in Immunol.* 10:478-482, 1998).

Secondary anchor residues that characterize the A2.1 motif have additionally been defined as disclosed herein. These are disclosed in Table II. Peptide binding to HLA-A2.1 molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Representative peptide epitopes that comprise an A2.1 motif are set forth on the attached Table VII. The A2.1 motifs comprising the primary anchor residues V, A, T, or Q at position 2 and L, I, V, A, or T at the C-terminal position are those most particularly relevant to the invention claimed herein.

#### **IV.D.12 HLA-A3 motif**

The allele-specific HLA-A3 motif is characterized by the presence in peptide ligands of L, M, V, I, S, A, T, F, C, G, or D as a primary anchor residue at position 2, and the presence of K, Y, R, H, F, or A as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA-A3 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Representative peptide epitopes that comprise the A3 motif are set forth on the attached Table XVI. Those peptide epitopes that also comprise the A3 supermotif are also listed in Table IX.

#### **IV.D.13. HLA-A11 motif**

The allele-specific HLA-A11 motif is characterized by the presence in peptide ligands of V, T, M, L, I, S, A, G, N, C, D, or F as a primary anchor residue in position 2, and K, R, Y, or H as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA-A11 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Representative peptide epitopes that comprise the A11 motif are set forth on the attached Table XVII; peptide epitopes comprising the A3 allele-specific motif are also present in this Table because of the extensive overlap between the A3 and A11 motif primary anchor specificities. Further, those peptide epitopes that comprise the A3 supermotif are also listed in Table IX.

#### IV.D.14. HLA-A24 motif

The allele-specific HLA-A24 motif is characterized by the presence in peptide ligands of Y, F, W, or M as a primary anchor residue in position 2, and F, L, I, or W as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA-A24 molecules can be modulated by substitutions at primary and/or secondary anchor positions; preferably choosing respective residues specified for the motif.

Representative peptide epitopes that comprise the A24 motif are set forth on the attached Table XVIII. These epitopes are also listed in Table X, which sets forth HLA-A24-supermotif-bearing peptide epitopes.

#### Motifs Indicative of Class II HTL Inducing Peptide Epitopes

The primary and secondary anchor residues of the HLA class II peptide epitope supermotifs and motifs delineated below are summarized in Table III.

#### IV.D.15. HLA DR-1-4-7 supermotif

Motifs have also been identified for peptides that bind to three common HLA class II allele-specific HLA molecules: HLA DRB1\*0401, DRB1\*0101, and DRB1\*0701. Collectively, the common residues from these motifs delineate the HLA DR-1-4-7 supermotif. Peptides that bind to these DR molecules carry a supermotif characterized by a large aromatic or hydrophobic residue (Y, F, W, L, I, V, or M) as a primary anchor residue in position 1, and a small, non-charged residue (S, T, C, A, P, V, I, L, or M) as a primary anchor residue in position 6 of the epitope. Allele-specific secondary effects and secondary anchors for each of these HLA types have also been identified. These are set forth in Table III. Peptide binding to HLA-DR4, DR1, and/or DR7 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Conserved peptide epitopes (*i.e.* 75% conservancy in the 20 HBV strains used for the analysis), corresponding to a nine residue core comprising the DR-1-4-7 supermotif

(wherein position 1 of the motif is at position 1 of the nine residue core) are set forth in Table XIXa (*see, e.g.*, Madden, *Annu. Rev. Immunol.* 13:587-622, 1995). Respective exemplary peptide epitopes of 15 amino acid residues in length, each of which comprise a conserved nine residue core, are also shown in section “a” of the Table. Cross-reactive binding data for the exemplary 15-residue supermotif-bearing peptides denoted by a peptide number are shown in Table XIXb.

#### IV.D.16. HLA DR3 motifs

Two alternative motifs (*i.e.*, submotifs) characterize peptide epitopes that bind to HLA-DR3 molecules. In the first motif (submotif DR3A) a large, hydrophobic residue (L, I, V, M, F, or Y) is present in anchor position 1, and D is present as an anchor at position 4, towards the carboxyl terminus of the epitope.

The alternative DR3 submotif provides for lack of the large, hydrophobic residue at anchor position 1, and/or lack of the negatively charged or amide-like anchor residue at position 4, by the presence of a positive charge at position 6 towards the carboxyl terminus of the epitope. Thus, for the alternative allele-specific DR3 motif (submotif DR3B): L, I, V, M, F, Y, A, or Y is present at anchor position 1; D, N, Q, E, S, or T is present at anchor position 4; and K, R, or H is present at anchor position 6. Peptide binding to HLA-DR3 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Conserved peptide epitopes (*i.e.*, sequences that are 75% conserved in the 20 HBV strains used for the analysis), corresponding to a nine residue core comprising the DR3A submotif (wherein position 1 of the motif is at position 1 of the nine residue core) set forth in Table XXa. Respective exemplary peptide epitopes of 15 amino acid residues in length, each of which comprise a conserved nine residue core, are also shown in section “a” of the Table. Table XXb shows binding data of the exemplary DR3 submotif A-bearing peptides denoted by a peptide number.

Conserved peptide epitopes (*i.e.*, 75% conservancy in the 20 HBV strains used for the analysis), corresponding to a nine residue core comprising the DR3B submotif and respective exemplary 15-mer peptides comprising the DR3 submotif-B epitope are set forth in Table XXc. Table XXd shows binding data of the exemplary DR3 submotif B-bearing peptides denoted by a peptide number.

Each of the HLA class I or class II peptide epitopes set out in the Tables herein are deemed singly to be an inventive aspect of this application. Further, it is also an

inventive aspect of this application that each peptide epitope may be used in combination with any other peptide epitope.

#### **IV.E. Enhancing Population Coverage of the Vaccine**

Vaccines that have broad population coverage are preferred because they are more commercially viable and generally applicable to the most people. Broad population coverage can be obtained using the peptides of the invention (and nucleic acid compositions that encode such peptides) through selecting peptide epitopes that bind to HLA alleles which, when considered in total, are present in most of the population. Table XXI lists the overall frequencies of the HLA class I supertypes in various ethnicities (Table XXIa) and the combined population coverage achieved by the A2-, A3-, and B7-supertypes (Table XXIb). The A2-, A3-, and B7 supertypes are each present on the average of over 40% in each of these five major ethnic groups. Coverage in excess of 80% is achieved with a combination of these supermotifs. These results suggest that effective and non-ethnically biased population coverage is achieved upon use of a limited number of cross-reactive peptides. Although the population coverage reached with these three main peptide specificities is high, coverage can be expanded to reach 95% population coverage and above, and more easily achieve truly multispecific responses upon use of additional supermotif or allele-specific motif bearing peptides.

The B44-, A1-, and A24-supertypes are present, on average, in a range from 25% to 40% of these major ethnic populations (Table XXIa). While less prevalent overall, the B27-, B58-, and B62 supertypes are each present with a frequency >25% in at least one major ethnic group (Table XXIa). Table XXIb summarizes the estimated combined prevalence in five major ethnic groups of HLA supertypes that have been identified. The incremental coverage obtained by the inclusion of A1-, A24-, and B44-supertypes to the A2, A3, and B7 coverage, or all of the supertypes described herein is shown. By including epitopes from the six most frequent supertypes, an average population coverage of 99% is obtained for five major ethnic groups.

The data presented herein, together with the previous definition of the A2-, A3-, and B7-supertypes, indicates that all antigens, with the possible exception of A29, B8, and B46, can be classified into a total of nine HLA supertypes. Focusing on the six most common supertypes affords population coverage greater than 98% for all major ethnic populations.

#### IV.F. Immune Response Stimulating Peptide Analogs

Although peptides with suitable cross-reactivity among all alleles of a superfamily are identified by the screening procedures described above, cross-reactivity is not always complete and in such cases procedures to further increase cross-reactivity of peptides can be useful; such procedures can also be used to modify other properties of the peptides. Having established the general rules that govern cross-reactivity of peptides for HLA alleles within a given motif or supermotif, modification (*i.e.*, analoging) of the structure of peptides of particular interest in order to achieve broader (or otherwise modified) HLA binding capacity can be performed. More specifically, peptides which exhibit the broadest cross-reactivity patterns, (both amongst the known T cell epitopes, as well as the more extended set of peptides that contain the appropriate supermotifs), can be produced in accordance with the teachings herein.

The strategy employed utilizes the motifs or supermotifs which correlate with binding to certain HLA molecules. The motifs or supermotifs are defined by having primary anchors, though secondary anchors can also be modified. Analog peptides can be created by substituting amino acids residues at primary anchor, secondary anchor, or at primary and secondary anchor positions. Generally, analogs are made for peptides that already bear a motif or supermotif. Preferred secondary anchor residues of supermotifs and motifs that have been defined for HLA class I and class II binding peptides are shown in Tables II and III, respectively.

For a number of the motifs or supermotifs in accordance with the invention, residues are defined which are deleterious to binding to allele-specific HLA molecules or members of HLA supertypes that bind to the respective motif or supermotif (Tables II and III). Accordingly, removal of residues that are detrimental to binding can be performed in accordance with the present invention. For example, in the case of the A3 supertype, when all peptides that have such deleterious residues are removed from the population of analyzed peptides, the incidence of cross-reactivity increases from 22% to 37% (see, e.g., Sidney, J. *et al.*, *Hu. Immunol.* 45:79, 1996). Thus, one strategy to improve the cross-reactivity of peptides within a given supermotif is simply to delete one or more of the deleterious residues present within a peptide and substitute a small "neutral" residue such as Ala (that may not influence T cell recognition of the peptide). An enhanced likelihood of cross-reactivity is expected if, together with elimination of detrimental residues within a peptide, residues associated with high affinity binding to multiple alleles within a superfamily are inserted.

To ensure that an analog peptide, when used as a vaccine, actually elicits a CTL response to the native epitope *in vivo* (or, in the case of class II epitopes, elicits helper T cells that cross-react with the wild type peptides), the analog peptide may be used to immunize T cells *in vitro* from individuals of the appropriate HLA allele. Thereafter, the immunized cells' capacity to induce lysis of wild type peptide sensitized target cells is evaluated. It will be desirable to use as antigen presenting cells, cells that have been either infected, or transfected with the appropriate genes, or, in the case of class II epitopes only, cells that have been pulsed with whole protein antigens, to establish whether endogenously produced antigen is also recognized by the relevant T cells.

Another embodiment of the invention to ensure adequate numbers of cross-reactive cellular binders is to create analogs of weak binding peptides. Class I peptides exhibiting binding affinities of 500-50000nM, and carrying an acceptable but suboptimal primary anchor residue at one or both positions can be "fixed" by substituting preferred anchor residues in accordance with the respective supertype. The analog peptides can then be tested for crossbinding activity.

Another embodiment for generating effective peptide analogs involves the substitution of residues that have an adverse impact on peptide stability or solubility in a liquid environment. This substitution may occur at any position of the peptide epitope. For example, a cysteine (C) can be substituted out in favor of  $\alpha$ -amino butyric acid. Due to its chemical nature, cysteine has the propensity to form disulfide bridges and sufficiently alter the peptide structurally so as to reduce binding capacity. Substituting  $\alpha$ -amino butyric acid for C not only alleviates this problem, but actually improves binding and crossbinding capability in certain instances (Review: A. Sette *et al.*, In: Persistent Viral Infections, Eds. R. Ahmed and I. Chen, John Wiley & Sons, England, 1999).

Substitution of cysteine with  $\alpha$ -amino butyric acid may occur at any residue of a peptide epitope, i.e. at either anchor or non-anchor positions.

In general, CTL and HTL responses are not directed against all possible epitopes. Rather, they are restricted to a few immunodominant determinants (Zinkernagel, *et al.*, *Adv. Immunol.* 27:5159, 1979; Bennink, *et al.*, *J. Exp. Med.* 168:1935-1939, 1988; Rawle, *et al.*, *J. Immunol.* 146:3977-3984, 1991). It has been recognized that immunodominance (Benacerraf, *et al.*, *Science* 175:273-279, 1972) could be explained by either the ability of a given epitope to selectively bind a particular HLA protein (determinant selection theory) (Vitiello, *et al.*, *J. Immunol.* 131:1635, 1983); Rosenthal, *et al.*, *Nature* 267:156-158, 1977), or being selectively recognized by the existing TCR (T cell receptor)

specificities (repertoire theory) (Klein, J., IMMUNOLOGY, THE SCIENCE OF SELF/NONSELF DISCRIMINATION, John Wiley & Sons, New York, pp. 270-310, 1982). It has been demonstrated that additional factors, mostly linked to processing events, can also play a key role in dictating, beyond strict immunogenicity, which of the many potential determinants will be presented as immunodominant (Sercarz, *et al.*, *Annu. Rev. Immunol.* 11:729-766, 1993).

The concept of dominance and subdominance is relevant to immunotherapy of both infectious diseases and cancer. For example, in the course of chronic viral disease, recruitment of subdominant epitopes can be important for successful clearance of the infection, especially if dominant CTL or HTL specificities have been inactivated by functional tolerance, suppression, mutation of viruses and other mechanisms (Franco, *et al.*, *Curr. Opin. Immunol.* 7:524-531, (1995)). In the case of cancer and tumor antigens, CTLs recognizing at least some of the highest binding affinity peptides might be functionally inactivated. Lower binding affinity peptides are preferentially recognized at these times.

In particular, it has been noted that a significant number of epitopes derived from known non-viral tumor associated antigens (TAA) bind HLA class I with intermediate affinity ( $IC_{50}$  in the 50-500 nM range). For example, it has been found that 8 of 15 known TAA peptides recognized by tumor infiltrating lymphocytes (TIL) or CTL bound in the 50-500 nM range. (These data are in contrast with estimates that 90% of known viral antigens that were recognized as peptides bound HLA with  $IC_{50}$  of 50 nM or less, while only approximately 10% bound in the 50-500 nM range (Sette, *et al.*, *J. Immunol.*, 153:558-5592 (1994)). In the cancer setting this phenomenon is probably due to elimination, or functional inhibition of the CTL recognizing several of the highest binding peptides, presumably because of T cell tolerization events.

Without intending to be bound by theory, it is believed that because T cells to dominant epitopes may have been clonally deleted, selecting subdominant epitopes may allow extant T cells to be recruited, which will then lead to a therapeutic response. However, the binding of HLA molecules to subdominant epitopes is often less vigorous than to dominant ones. Accordingly, there is a need to be able to modulate the binding affinity of particular immunogenic epitopes for one or more HLA molecules, and thereby to modulate the immune response elicited by the peptide. Thus, a need exists to prepare analog peptides which elicit a more vigorous response. This ability would greatly enhance the usefulness of peptide-based vaccines and therapeutic agents.



Representative analog peptides are set forth in Table XXII. The Table indicates the length and sequence of the analog peptide as well as the motif or supermotif, if appropriate. The information in the "Fixed Nomenclature" column indicates the residues substituted at the indicated position numbers for the respective analog.

5

#### **IV.G. Computer Screening of Protein Sequences from Disease-Related Antigens for Supermotif or Motif Containing Peptides**

Computer programs that allow the rapid screening of protein sequences for the occurrence of the subject supermotifs or motifs are encompassed by the present invention; as are programs that permit the generation of analog peptides. These programs are implemented to analyze any identified amino acid sequence or operate on an unknown sequence and simultaneously determine the sequence and identify motif-bearing epitopes thereof; analogs can be simultaneously determined as well. Generally, the identified sequences will be from a pathogenic organism or a tumor-associated peptide. For example, the target molecules considered herein include all of the HBV proteins (e.g. surface, core, polymerase, and X).

In cases where the sequence of multiple variants of the same target protein are available, peptides may also be selected on the basis of their conservancy. A presently preferred criterion for conservancy defines that the entire sequence of a peptide be totally conserved in 75% of the sequences evaluated for a specific protein; this definition of conservancy has been employed herein.

It is important that the selection criteria utilized for prediction of peptide binding are as accurate as possible, to correlate most efficiently with actual binding. Prediction of peptides that bind, for example, to HLA-A\*0201, on the basis of the presence of the appropriate primary anchors, is positive at about a 30% rate (Ruppert, J. *et al. Cell* 74:929, 1993). However, by analyzing an extensive peptide-HLA binding database, the present inventors have developed a number of allele specific polynomial algorithms that dramatically increase the predictive value over identification on the basis of the presence of primary anchor residues alone. These algorithms take into account not only the presence or absence of the correct primary anchors, but also consider the positive or deleterious presence of secondary anchor residues (to account for the impact of different amino acids at different positions). The algorithms are essentially based on the premise

that the overall affinity (or  $\Delta G$ ) of peptide-HLA interactions can be approximated as a linear polynomial function of the type:

$$\Delta G = a_{1i} \times a_{2i} \times a_{3i} \dots \times a_{ni}$$

where  $a_{ij}$  is a coefficient that represents the effect of the presence of a given amino acid ( $j$ ) at a given position ( $i$ ) along the sequence of a peptide of  $n$  amino acids. An important assumption of this method is that the effects at each position are essentially independent of each other. This assumption is justified by studies that demonstrated that peptides are bound to HLA molecules and recognized by T cells in essentially an extended conformation. Derivation of specific algorithm coefficients has been described in Gulukota *et al.* (Gulukota, K. *et al.*, *J.Mol.Biol.* 267:1258, 1997).

Additional methods to identify preferred peptide sequences, which also make use of specific motifs, include the use of neural networks and molecular modeling programs (Gulukota, K. *et al.*, *J.Mol.Biol.* 267:1258, 1997; Milik *et al.*, *Nature Biotechnology* 16:753, 1998; Altuvia *et al.*, *Hum. Immunol.* 58:1, 1997; Altuvia *et al.*, *J. Mol. Biol.* 249:244, 1995; Buus, S. *Curr. Opin. Immunol.* 11:209-213, 1999; Brusic, V. *et al.*, *Bioinformatics* 14:121-130, 1998).

For example, it has been shown that in sets of A\*0201 motif peptides, 69% of the peptides containing at least one preferred secondary anchor residue while avoiding the presence of any deleterious secondary anchor residues, will bind A\*0201 with an  $IC_{50}$  less than 500 nM (Ruppert, J. *et al.* *Cell* 74:929, 1993). These algorithms are also flexible in that cut-off scores may be adjusted to select sets of peptides with greater or lower predicted binding properties, as desired.

In utilizing computer screening to identify peptide epitopes, all protein sequence or translated sequence may be analyzed using software developed to search for motifs, for example the "FINDPATTERNS" program (Devereux, *et al.* *Nucl. Acids Res.* 12:387-395, 1984) or MotifSearch 1.4 software program (D. Brown, San Diego, CA) to identify potential peptide sequences containing appropriate HLA binding motifs. As appreciated by one of ordinary skill in the art a large array of software and hardware options are available which can be employed to implement the motifs of the invention relative to known or unknown peptide sequences. The identified peptides will then be scored using customized polynomial algorithms to predict their capacity to bind specific HLA class I or class II alleles.

In accordance with the procedures described above, HBV peptides and analogs thereof that are able to bind HLA supertype groups or allele-specific HLA molecules have been identified (Tables VII-XX; Table XXII).

#### 5 **IV.H. Preparation of Peptide Epitopes**

Peptides in accordance with the invention can be prepared synthetically, by recombinant DNA technology, or from natural sources such as native tumors or pathogenic organisms. Peptide epitopes may be synthesized individually or as polyepitopic peptides. Although the peptide will preferably be substantially free of other  
 10 naturally occurring host cell proteins and fragments thereof, in some embodiments the peptides may be synthetically conjugated to native fragments or particles.

The peptides in accordance with the invention can be a variety of lengths, and either in their neutral (uncharged) forms or in forms which are salts. The peptides in accordance with the invention are either free of modifications such as glycosylation, side  
 15 chain oxidation, or phosphorylation; or they contain these modifications, subject to the condition that modifications do not destroy the biological activity of the peptides as described herein.

Desirably, the peptide will be as small as possible while still maintaining substantially all of the biological activity of the large peptide. When possible, it may be  
 20 desirable to optimize HLA class I binding peptides of the invention to a length of about 8 to about 13 amino acid residues, preferably 9 to 10. HLA class II binding peptides may be optimized to a length of about 6 to about 25 amino acids in length, preferably to between about 13 and about 20 residues. Preferably, the peptides are commensurate in size with endogenously processed pathogen-derived peptides or tumor cell peptides that  
 25 are bound to the relevant HLA molecules. Moreover, the identification and preparation of peptides of other lengths can be carried out using the techniques described herein (e.g., the disclosures regarding primary and secondary anchor positions). However, it is also preferred to identify a larger region of a native peptide that encompasses one and preferably two or more epitopes in accordance with the invention. This sequence is  
 30 selected on the basis that it contains the greatest number of epitopes per amino acid length. It is to be appreciated that epitopes can be present in a frame-shifted manner, e.g. a 10 amino acid long peptide could contain two 9 amino acid long epitopes and one 10 amino acid long epitope; each epitope can be exposed and bound by an HLA molecule upon administration of a plurality of such peptides. This larger, preferably multi-epitopic,

peptide can then be generated synthetically, recombinantly, or via cleavage from the native source.

The peptides of the invention can be prepared in a wide variety of ways. For the preferred relatively short size, the peptides can be synthesized in solution or on a solid support in accordance with conventional techniques. Various automatic synthesizers are commercially available and can be used in accordance with known protocols. See, for example, Stewart & Young, SOLID PHASE PEPTIDE SYNTHESIS, 2D. ED., Pierce Chemical Co. (1984). Further, individual peptides may be joined using chemical ligation to produce larger peptides.

Alternatively, recombinant DNA technology may be employed wherein a nucleotide sequence which encodes an immunogenic peptide of interest is inserted into an expression vector, transformed or transfected into an appropriate host cell and cultivated under conditions suitable for expression. These procedures are generally known in the art, as described generally in Sambrook *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989). Thus, recombinant polypeptides which comprise one or more peptide sequences of the invention can be used to present the appropriate T cell epitope.

As the nucleotide coding sequence for peptides of the preferred lengths contemplated herein can be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci, *et al.*, *J. Am. Chem. Soc.* 103:3185 (1981) modification can be made simply by substituting the appropriate and desired nucleic acid base(s) for those that encode the native peptide sequence. The coding sequence can then be provided with appropriate linkers and ligated into expression vectors commonly available in the art, and the vectors used to transform suitable hosts to produce the desired fusion protein. A number of such vectors and suitable host systems are now available. For expression of the fusion proteins, the coding sequence will be provided with operably linked start and stop codons, promoter and terminator regions and usually a replication system to provide an expression vector for expression in the desired cellular host. For example, promoter sequences compatible with bacterial hosts are provided in plasmids containing convenient restriction sites for insertion of the desired coding sequence. The resulting expression vectors are transformed into suitable bacterial hosts. Of course, yeast, insect or mammalian cell hosts may also be used, employing suitable vectors and control sequences.

#### IV.I. Assays to Detect T-Cell Responses

Once HLA binding peptides are identified, they can be tested for the ability to elicit a T-cell response. The preparation and evaluation of motif-bearing peptides are described in PCT publications WO 94/20127 and WO 94/03205. Briefly, peptides comprising epitopes from a particular antigen are synthesized and tested for their ability to bind to the appropriate HLA proteins in assays using, for example, purified HLA class I molecules and radioiodinated peptides and/or cells expressing empty class I molecules (which lack peptide in their receptor) by, for instance, immunofluorescent staining and flow microfluorimetry, peptide-dependent class I assembly assays, and inhibition of CTL recognition by peptide competition. Those peptides that bind to the class I molecule are further evaluated for their ability to serve as targets for CTLs derived from infected or immunized individuals, as well as for their capacity to induce primary *in vitro* or *in vivo* CTL responses that can give rise to CTL populations capable of reacting with selected target cells associated with a disease. Corresponding assays are used for evaluation of HLA class II binding peptides.

Conventional assays utilized to detect CTL responses include proliferation assays, lymphokine secretion assays, direct cytotoxicity assays, and limiting dilution assays. For example, antigen-presenting cells that have been incubated with a peptide can be assayed for the ability to induce CTL responses in responder cell populations. Antigen-presenting cells can be normal cells such as peripheral blood mononuclear cells or dendritic cells. Alternatively, mutant mammalian cell lines that are deficient in their ability to load class I molecules with internally processed peptides and that have been transfected with the appropriate human class I gene may be used to test for the capacity of the peptide to induce *in vitro* primary CTL responses.

Peripheral blood lymphocytes may be used as the responder cell source of CTL precursors. The appropriate antigen-presenting cells are incubated with peptide and the peptide-loaded antigen-presenting cells are then incubated with the responder cell population under optimized culture conditions. Positive CTL activation can be determined by assaying the culture for the presence of CTLs that kill radio-labeled target cells, both specific peptide-pulsed targets as well as target cells expressing endogenously processed forms of the HBV antigen from which the peptide sequence was derived.

More recently, a method has also been devised which allows direct quantification of antigen-specific T cells by staining with Fluorescein-labelled HLA tetrameric complexes (Altman, J. D. *et al.*, *Proc. Natl. Acad. Sci. USA* 90:10330, 1993; Altman, J.

D. *et al.*, *Science* 274:94, 1996). Other relatively recent technical developments include staining for intracellular lymphokines, and interferon release assays or ELISPOT assays. Tetramer staining, intracellular lymphokine staining and ELISPOT assays all appear to be at least 10-fold more sensitive than more conventional assays (Lalvani, A. *et al.*, *J. Exp. Med.* 186:859, 1997; Dunbar, P. R. *et al.*, *Curr. Biol.* 8:413, 1998; Murali-Krishna, K. *et al.*, *Immunity* 8:177, 1998).

HTL activation may also be assessed using such techniques as T cell proliferation and secretion of lymphokines, e.g. IL-2.

Alternatively, immunization of HLA transgenic mice can be used to determine immunogenicity of peptide epitopes. Several transgenic mouse models including mice with human A2.1, A11, and B7 alleles have been characterized and others (e.g., transgenic mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed. Additional transgenic mouse models with other HLA alleles may be generated as necessary. Mice may be immunized with peptides emulsified in Incomplete Freund's Adjuvant and the resulting T cells tested for their capacity to recognize peptide-pulsed target cells and target cells transfected with appropriate genes. CTL responses may be analyzed using cytotoxicity assays described above. Similarly, HTL responses may be analyzed using such assays as T cell proliferation or secretion of lymphokines.

Immunogenic peptide epitopes are set out in Table XXIII.

#### **IV.J. Use of Peptide Epitopes as Diagnostic Agents and for Evaluating Immune Responses**

HLA class I and class II binding peptides as described herein can be used, in one embodiment of the invention, as reagents to evaluate an immune response. The immune response to be evaluated may be induced by using as an immunogen any agent that would potentially result in the production of antigen-specific CTLs or HTLs to the peptide epitope(s) to be employed as the reagent. The peptide reagent is not used as the immunogen.

For example, a peptide of the invention may be used in a tetramer staining assay to assess peripheral blood mononuclear cells for the presence of antigen-specific CTLs following exposure to a pathogen or immunogen. The HLA-tetrameric complex is used to directly visualize antigen-specific CTLs (*see, e.g.*, Ogg *et al.* *Science* 279:2103-2106, 1998; and Altman *et al.* *Science* 174:94-96, 1996) and determine the frequency of the

antigen-specific CTL population in a sample of peripheral blood mononuclear cells. A tetramer reagent using a peptide of the invention may be generated as follows: A peptide that binds to an allele-specific HLA molecules, or supertype molecules, is refolded in the presence of the corresponding HLA heavy chain and  $\beta_2$ -microglobulin to generate a trimolecular complex. The complex is biotinylated at the carboxyl terminal end of the heavy chain at a site that was previously engineered into the protein. Tetramer formation is then induced by the addition of streptavidin. By means of fluorescently labeled streptavidin, the tetramer can be used to stain antigen-specific cells. The cells may then be identified, for example, by flow cytometry. Such an analysis may be used for diagnostic or prognostic purposes.

Peptides of the invention may also be used as reagents to evaluate immune recall responses. (*see, e.g., Berton et al. J. Clin. Invest.* 100:503-513, 1997 and *Penna et al. J. Exp. Med.* 174:1565-1570, 1991.) For example, patient PBC samples from individuals with acute hepatitis B or who have recently recovered from acute hepatitis B may be analyzed for the presence of HBV antigen-specific CTLs using HBV-specific peptides. A blood sample containing mononuclear cells may be evaluated by cultivating the PBCs and stimulating the cells with a peptide of the invention. After an appropriate cultivation period, the expanded cell population may be analyzed for cytotoxic activity.

The peptides may also be used as reagents to evaluate the efficacy of a vaccine. PBMCs obtained from a patient vaccinated with an immunogen may be analyzed using, for example, either of the methods described above. A patient is HLA typed, and appropriate peptide reagents that recognize allele-specific molecules present in that patient may be selected for the analysis. The immunogenicity of the vaccine will be indicated by the presence of HBV epitope-specific CTLs in the PBMC sample.

The peptides of the invention may also be used to make antibodies using techniques well known in the art (*see, e.g., CURRENT PROTOCOLS IN IMMUNOLOGY*, Wiley/Greene, NY; and *Antibodies A Laboratory Manual*, Harlow and Lane, Cold Spring Harbor Laboratory Press, 1989). Such antibodies may be useful as reagents to diagnose HBV infection.

#### IV.K. Vaccine Compositions

Vaccines that contain an immunogenically effective amount of one or more peptides as described herein are a further embodiment of the invention. Once

appropriately immunogenic epitopes have been defined, they can be sorted and delivered by various means, herein referred to as “vaccine” compositions. Such vaccine compositions can include, for example, lipopeptides (Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341, 1995), peptides compositions encapsulated in poly(DL-lactide-co-glycolide) (PLG) microspheres (see, e.g., Eldridge, *et al. Molec. Immunol.* 28:287-294, 1991; *Alonso et al. Vaccine* 12:299-306, 1994; Jones *et al. Vaccine* 13:675-681, 1995), peptide compositions encapsulated in immune stimulating complexes (ISCOMS) (see, e.g., Takahashi *et al. Nature* 344:873-875, 1990; Hu *et al. Clin Exp Immunol.* 113:235-243, 1998), multiple antigen peptide systems (MAPs) (see e.g., Tam, J. P., *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413, 1988; Tam, J.P., *J. Immunol. Methods* 196:17-32, 1996), viral delivery vectors (Perkus, M. E. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 379, 1996; Chakrabarti, S. *et al.*, *Nature* 320:535, 1986; Hu, S. L. *et al.*, *Nature* 320:537, 1986; Kieny, M.-P. *et al.*, *AIDS Bio/Technology* 4:790, 1986; Top, F. H. *et al.*, *J. Infect. Dis.* 124:148, 1971; Chanda, P. K. *et al.*, *Virology* 175:535, 1990), particles of viral or synthetic origin (Kofler, N. *et al.*, *J. Immunol. Methods.* 192:25, 1996; Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993; Falo, L. D., Jr. *et al.*, *Nature Med.* 7:649, 1995), adjuvants (Warren, H. S., Vogel, F. R., and Chedid, L. A. *Annu. Rev. Immunol.* 4:369, 1986; Gupta, R. K. *et al.*, *Vaccine* 11:293, 1993), liposomes (Reddy, R. *et al.*, *J. Immunol.* 148:1585, 1992; Rock, K. L., *Immunol. Today* 17:131, 1996), or, naked or particle absorbed cDNA (Ulmer, J. B. *et al.*, *Science* 259:1745, 1993; Robinson, H. L., Hunt, L. A., and Webster, R. G., *Vaccine* 11:957, 1993; Shiver, J. W. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., *Annu. Rev. Immunol.* 12:923, 1994 and Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993). Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Furthermore, vaccines in accordance with the invention encompass compositions of one or more of the claimed peptide(s) that can be introduced into a host, including humans, linked to its own carrier, or as a homopolymer or heteropolymer of active peptide units. Such a polymer has the advantage of increased immunological reaction and, where different peptides are used to make up the polymer, the additional ability to induce antibodies and/or CTLs that react with different antigenic determinants of the pathogenic organism or tumor-related peptide targetted for an immune response.



Furthermore, useful carriers that can be used with vaccines of the invention are well known in the art, and include, *e.g.*, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza, hepatitis B virus core protein, and the like. The vaccines can contain a

5 physiologically tolerable (i.e., acceptable) diluent such as water, or saline, preferably phosphate buffered saline. The vaccines also typically include an adjuvant. Adjuvants such as incomplete Freund's adjuvant, aluminum phosphate, aluminum hydroxide, or alum are examples of materials well known in the art. Additionally, as disclosed herein, CTL responses can be primed by conjugating peptides of the invention to lipids, such as

10 tripalmitoyl-S-glycerylcysteinylserine (P<sub>3</sub>CSS).

As disclosed in greater detail herein, upon immunization with a peptide composition in accordance with the invention, via injection, aerosol, oral, transdermal, transmucosal, intrapleural, intrathecal, or other suitable routes, the immune system of the host responds to the vaccine by producing large amounts of CTLs specific for the desired

15 antigen. Consequently, the host becomes at least partially immune to later infection, or at least partially resistant to developing an ongoing chronic infection, or derives at least some therapeutic benefit when the antigen was tumor-associated.

In some instances it may be desirable to combine the class I peptide vaccines of the invention with vaccines which induce or facilitate neutralizing antibody responses to

20 the target antigen of interest, particularly to viral envelope antigens. A preferred embodiment of such a composition comprises class I and class II epitopes in accordance with the invention. An alternative embodiment of such a composition comprises a class I and/or class II epitope in accordance with the invention, along with a PADRE™ (Epimmune, San Diego, CA) molecule (described in the related U.S.S.N. 08/485,218, which is a CIP of U.S.S.N. 08/305,871, now U.S. Patent Number 5,736,142, which is a

25 CIP of abandoned application U.S.S.N. 08/121,101.) Furthermore, any of these embodiments can be administered as a nucleic acid mediated modality.

For therapeutic or immunization purposes, the peptides of the invention can also be expressed by viral or bacterial vectors. Examples of expression vectors include

30 attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into an acutely or chronically infected host or into a non-infected host, the recombinant vaccinia virus expresses the immunogenic peptide, and

thereby elicits a host CTL and/or HTL response. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover, *et al. Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, *e.g.* adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein.

Antigenic peptides are used to elicit a CTL and/or HTL response *ex vivo*, as well. The resulting CTL or HTL cells, can be used to treat chronic infections, or tumors in patients that do not respond to other conventional forms of therapy, or will not respond to a therapeutic vaccine peptide or nucleic acid in accordance with the invention. *Ex vivo* CTL or HTL responses to a particular pathogen (infectious agent or tumor antigen) are induced by incubating in tissue culture the patient's CTL or HTL precursor cells together with a source of antigen-presenting cells (APC), such as dendritic cells, and the appropriate immunogenic peptide. After an appropriate incubation time (typically about 14 weeks), in which the precursor cells are activated, mature and expand into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cell (an infected cell or a tumor cell). Transfected dendritic cells may also be used as antigen presenting cells. Alternatively, dendritic cells are transfected, *e.g.*, with a minigene construct in accordance with the invention, in order to elicit immune responses. Minigenes will be discussed in greater detail in a following section.

DNA or RNA encoding one or more of the peptides of the invention can also be administered to a patient. This approach is described, for instance, in Wolff *et al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include “naked DNA”, facilitated (bupivacaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated (“gene gun”) delivery.

Preferably, the following principles are utilized when selecting an array of epitopes for inclusion in a polyepitopic composition, or for selecting epitopes to be included in a vaccine composition and/or to be encoded by a minigene. It is preferred that each of the following principles are balanced in order to make the selection.

1.) Epitopes are selected which, upon administration, mimic immune responses that have been observed to be correlated with HBV clearance. For HLA Class I this includes 3-4 epitopes that come from at least one antigen of HBV. In other words, it has been observed that in patients who spontaneously clear HBV, that they had generated an immune response to at least 3 epitopes on at least one HBV antigen. For HLA Class II a similar rationale is employed; again 3-4 epitopes are selected from at least one HBV antigen (see *e.g.*, Rosenberg *et al. Science* 278:1447-1450).

2.) Epitopes are selected that have the requisite binding affinity established to be correlated with immunogenicity: for HLA Class I an  $IC_{50}$  of 500 nM or less, or for Class II an  $IC_{50}$  of 1000 nM or less.

3.) Sufficient supermotif bearing peptides, or a sufficient array of allele-specific motif bearing peptides, are selected to give broad population coverage. For example, it is preferable to have at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess population coverage.

4.) When selecting epitopes from cancer-related antigens it is often preferred to select analogs. When selecting epitopes for infectious disease-related antigens it is often preferable to select native epitopes. Therefore, of particular relevance for infectious disease vaccines (but for cancer-related vaccines as well), are epitopes referred to as “nested epitopes.” Nested epitopes occur where at least two epitopes overlap in a given peptide sequence. A peptide comprising “transcendent nested epitopes” is a peptide that has both HLA class I and HLA class II epitopes in it.

When providing nested epitopes, it is preferable to provide a sequence that has the greatest number of epitopes per provided sequence. A limitation on this principle is to avoid providing a peptide that is any longer than the amino terminus of the amino terminal epitope and the carboxyl terminus of the carboxyl terminal epitope in the peptide. When providing a longer peptide sequence, such as a sequence comprising nested epitopes, it is important to screen the sequence in order to insure that it does not have pathological or other deleterious biological properties.

5.) When creating a minigene, as disclosed in greater detail in the following section, an objective is to generate the smallest peptide possible that encompasses the epitopes of interest. The principles employed are similar, if not the same as those employed when selecting a peptide comprising nested epitopes. Thus, upon determination of the nucleic acid sequence to be provided as a minigene, the peptide

encoded thereby is analyzed to determine whether any "junctional epitopes" have been created. A junctional epitope is an actual binding epitope, as predicted, e.g., by motif analysis. Junctional epitopes are to be avoided because the recipient may generate an immune response to that epitope. Of particular concern is a junctional epitope that is a  
 5 "dominant epitope." A dominant epitope may lead to such a zealous response that immune responses to other epitopes are diminished or suppressed.

#### IV.K.1. Minigene Vaccines

A growing body of experimental evidence demonstrates that a number of different  
 10 approaches are available which allow simultaneous delivery of multiple epitopes.

Nucleic acids encoding the peptides of the invention are a particularly useful embodiment of the invention. Epitopes for inclusion in a minigene are preferably selected according to the guidelines set forth in the previous section. A preferred means of administering nucleic acids encoding the peptides of the invention uses minigene constructs encoding a  
 15 peptide comprising one or multiple epitopes of the invention. The use of multi-epitope minigenes is described below and in, e.g. An, L. and Whitton, J. L., *J. Virol.* 71:2292, 1997; Thomson, S. A. *et al.*, *J. Immunol.* 157:822, 1996; Whitton, J. L. *et al.*, *J. Virol.* 67:348, 1993; Hanke, R. *et al.*, *Vaccine* 16:426, 1998. For example, a multi-epitope DNA plasmid encoding nine dominant HLA-A\*0201- and A11-restricted epitopes derived from  
 20 the polymerase, envelope, and core proteins of HBV and HIV, the PADRE™ universal helper T cell (HTL) epitope, and an endoplasmic reticulum-translocating signal sequence was engineered. Immunization of HLA transgenic mice with this plasmid construct resulted in strong CTL induction responses against the nine epitopes tested, similar to those observed with a lipopeptide of known immunogenicity in humans, and significantly  
 25 greater than immunization in oil-based adjuvants. Moreover, the immunogenicity of DNA-encoded epitopes *in vivo* correlated with the *in vitro* responses of specific CTL lines against target cells transfected with the DNA plasmid. Thus, these data show that the minigene served to both: 1.) generate a CTL response and 2.) that the induced CTLs recognized cells expressing the encoded epitopes.

30 For example, to create a DNA sequence encoding the selected epitopes (minigene) for expression in human cells, the amino acid sequences of the epitopes may be reverse translated. A human codon usage table can be used to guide the codon choice for each amino acid. These epitope-encoding DNA sequences may be directly adjoined, so that

when translated, a continuous polypeptide sequence is created. To optimize expression and/or immunogenicity, additional elements can be incorporated into the minigene design. Examples of amino acid sequences that could be reverse translated and included in the minigene sequence include: HLA class I epitopes, HLA class II epitopes, a ubiquitination signal sequence, and/or an endoplasmic reticulum targeting signal. In addition, HLA presentation of CTL and HTL epitopes may be improved by including synthetic (*e.g.* poly-alanine) or naturally-occurring flanking sequences adjacent to the CTL or HTL epitopes; these larger peptides comprising the epitope(s) are within the scope of the invention.

The minigene sequence may be converted to DNA by assembling oligonucleotides that encode the plus and minus strands of the minigene. Overlapping oligonucleotides (30-100 bases long) may be synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides can be joined, for example, using T4 DNA ligase. This synthetic minigene, encoding the epitope polypeptide, can then be cloned into a desired expression vector.

Standard regulatory sequences well known to those of skill in the art are preferably included in the vector to ensure expression in the target cells. Several vector elements are desirable: a promoter with a down-stream cloning site for minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (*e.g.* ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, *e.g.*, the human cytomegalovirus (hCMV) promoter. See, *e.g.*, U.S. Patent Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences.

Additional vector modifications may be desired to optimize minigene expression and immunogenicity. In some cases, introns are required for efficient gene expression, and one or more synthetic or naturally-occurring introns could be incorporated into the transcribed region of the minigene. The inclusion of mRNA stabilization sequences and sequences for replication in mammalian cells may also be considered for increasing minigene expression.

Once an expression vector is selected, the minigene is cloned into the polylinker region downstream of the promoter. This plasmid is transformed into an appropriate *E. coli* strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the minigene, as well as all other elements included in the vector, are

confirmed using restriction mapping and DNA sequence analysis. Bacterial cells harboring the correct plasmid can be stored as a master cell bank and a working cell bank.

In addition, immunostimulatory sequences (ISSs or CpGs) appear to play a role in the immunogenicity of DNA vaccines. These sequences may be included in the vector,  
5 outside the minigene coding sequence, if desired to enhance immunogenicity.

In some embodiments, a bi-cistronic expression vector which allows production of both the minigene-encoded epitopes and a second protein (included to enhance or decrease immunogenicity) can be used. Examples of proteins or polypeptides that could beneficially enhance the immune response if co-expressed include cytokines (*e.g.*, IL-2,  
10 IL-12, GM-CSF), cytokine-inducing molecules (*e.g.*, LeIF) or costimulatory molecules. Helper (HTL) epitopes can be joined to intracellular targeting signals and expressed separately from expressed CTL epitopes; this allows direction of the HTL epitopes to a cell compartment different than that of the CTL epitopes. If required, this could facilitate more efficient entry of HTL epitopes into the HLA class II pathway, thereby improving  
15 CTL induction. In contrast to HTL or CTL induction, specifically decreasing the immune response by co-expression of immunosuppressive molecules (*e.g.* TGF- $\beta$ ) may be beneficial in certain diseases).

Therapeutic quantities of plasmid DNA can be produced for example, by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are  
20 used to inoculate growth medium, and grown to saturation in shaker flasks or a bioreactor according to well known techniques. Plasmid DNA can be purified using standard bioseparation technologies such as solid phase anion-exchange resins supplied by QIAGEN, Inc. (Valencia, California). If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

Purified plasmid DNA can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized DNA in sterile phosphate-buffer saline (PBS). This approach, known as "naked DNA," is currently being used for intramuscular (IM) administration in clinical trials. To maximize the immunotherapeutic effects of minigene DNA vaccines, an alternative method for  
25 formulating purified plasmid DNA may be desirable. A variety of methods have been described, and new techniques may become available. Cationic lipids can also be used in the formulation (*see, e.g.*, as described by WO 93/24640; Mannino & Gould-Fogerite, *BioTechniques* 6(7): 682 (1988); U.S. Pat No. 5,279,833; WO 91/06309; and Felgner, *et*  
30

*al., Proc. Nat'l Acad. Sci. USA* 84:7413 (1987). In addition, glycolipids, fusogenic liposomes, peptides and compounds referred to collectively as protective, interactive, non-condensing compounds (PINC) could also be complexed to purified plasmid DNA to influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

Target cell sensitization can be used as a functional assay for expression and HLA class I presentation of minigene-encoded CTL epitopes. For example, the plasmid DNA is introduced into a mammalian cell line that is suitable as a target for standard CTL chromium release assays. The transfection method used will be dependent on the final formulation. Electroporation can be used for "naked" DNA, whereas cationic lipids allow direct *in vitro* transfection. A plasmid expressing green fluorescent protein (GFP) can be co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). These cells are then chromium-51 ( $^{51}\text{Cr}$ ) labeled and used as target cells for epitope-specific CTL lines; cytolysis, detected by  $^{51}\text{Cr}$  release, indicates both production of, and HLA presentation of, minigene-encoded CTL epitopes.

*In vivo* immunogenicity is a second approach for functional testing of minigene DNA formulations. Transgenic mice expressing appropriate human HLA proteins are immunized with the DNA product. The dose and route of administration are formulation dependent (*e.g.*, IM for DNA in PBS, intraperitoneal (IP) for lipid-complexed DNA). Twenty-one days after immunization, splenocytes are harvested and restimulated for 1 week in the presence of peptides encoding each epitope being tested. Thereafter, for CTL effector cells, assays are conducted for cytolysis of peptide-loaded,  $^{51}\text{Cr}$  labeled target cells using standard techniques. Lysis of target cells sensitized by HLA loading of peptides corresponding to minigene-encoded epitopes demonstrates DNA vaccine function for *in vivo* induction of CTLs. Immunogenicity of HTL epitopes is evaluated in transgenic mice in an analogous manner.

Alternatively, the nucleic acids can be administered using ballistic delivery as described, for instance, in U.S. Patent No. 5,204,253. Using this technique, particles comprised solely of DNA are administered. In a further alternative embodiment, DNA can be adhered to particles, such as gold particles.

#### IV.K.2. Combinations of CTL Peptides with Helper Peptides

The peptides of the present invention, or analogs thereof, which have immunostimulatory activity may be modified to provide desired attributes, such as improved serum half life, or to enhance immunogenicity.

For instance, the ability of the peptides to induce CTL activity can be enhanced by linking the peptide to a sequence which contains at least one epitope that is capable of inducing a T helper cell response. The use of T helper epitopes in conjunction with CTL epitopes to enhance immunogenicity is illustrated, for example, in U.S.S.N. 08/820360, U.S.S.N. 08/197,484, U.S.S.N. 08/464,234, U.S.S.N. 08/464,496, U.S.S.N. 08/464,031, abandoned U.S.S.N. 08/464,433, and U.S.S.N. 08/461,603.

Particularly preferred CTL epitope/HTL epitope conjugates are linked by a spacer molecule. The spacer is typically comprised of relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological conditions. The spacers are typically selected from, *e.g.*, Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer need not be comprised of the same residues and thus may be a hetero- or homo-oligomer. When present, the spacer will usually be at least one or two residues, more usually three to six residues. Alternatively, the CTL peptide may be linked to the T helper peptide without a spacer.

The CTL peptide epitope may be linked to the HTL peptide epitope either directly or via a spacer either at the amino or carboxy terminus of the CTL peptide. The amino terminus of either the CTL epitope or the HTL peptide may be acylated. The HTL peptide epitopes used in the invention can be modified in the same manner as CTL peptides. For instance, they may be modified to include D-amino acids or be conjugated to other molecules such as lipids, proteins, sugars and the like. Exemplary T helper peptides include tetanus toxoid 830-843, influenza 307-319, and malarial circumsporozoite 382-398 and 378-389.

In certain embodiments, the T helper peptide is one that is recognized by T helper cells present in the majority of the population. This can be accomplished by selecting amino acid sequences that bind to many, most, or all of the HLA class II molecules. These are known as "loosely HLA-restricted" or "promiscuous" T helper sequences. Examples of amino acid sequences that are promiscuous include sequences from antigens such as tetanus toxoid at positions 830-843 (QYIKANSKFIGITE), *Plasmodium falciparum* CS protein at positions 378-398 (DIEKKIAKMEKASSVFNVVNS), and



Streptococcus 18kD protein at positions 116 (GAVDSILGGVATYGAA). Other examples include peptides bearing a DR 1-4-7 supermotif, or either of the DR3 motifs.

Alternatively, it is possible to prepare synthetic peptides capable of stimulating T helper lymphocytes, in a loosely HLA-restricted fashion, using amino acid sequences not found in nature (*see, e.g.*, PCT publication WO 95/07707). These synthetic compounds called Pan-DR-binding epitopes (*e.g.*, PADRE™, Epimmune, Inc., San Diego, CA) are designed on the basis of their binding activity to most HLA-DR (human HLA class II) molecules. For instance, a pan-DR-binding epitope peptide having the formula: aKXVWANTLKAAa, where “X” is either cyclohexylalanine, phenylalanine, or tyrosine, and a is either D-alanine or L-alanine, has been found to bind to most HLA-DR alleles, and to stimulate the response of T helper lymphocytes from most individuals, regardless of their HLA type.

HTL peptide epitopes can also be modified to alter their biological properties. For example, peptides comprising HTL epitopes can contain D-amino acids to increase their resistance to proteases and thus extend their serum half-life. Also, the epitope peptides of the invention can be conjugated to other molecules such as lipids, proteins or sugars, or any other synthetic compounds, to increase their biological activity. Specifically, the T helper peptide can be conjugated to one or more palmitic acid chains at either the amino or carboxyl termini.

In some embodiments it may be desirable to include in the pharmaceutical compositions of the invention at least one component which primes cytotoxic T lymphocytes. Lipids have been identified as agents capable of priming CTL *in vivo* against viral antigens. For example, palmitic acid residues can be attached to the  $\epsilon$ - and  $\alpha$ -amino groups of a lysine residue and then linked, *e.g.*, via one or more linking residues such as Gly, Gly-Gly-, Ser, Ser-Ser, or the like, to an immunogenic peptide. The lipidated peptide can then be administered either directly in a micelle or particle, incorporated into a liposome, or emulsified in an adjuvant, *e.g.*, incomplete Freund's adjuvant. In a preferred embodiment, a particularly effective immunogenic comprises palmitic acid attached to  $\epsilon$ - and  $\alpha$ - amino groups of Lys, which is attached via linkage, *e.g.*, Ser-Ser, to the amino terminus of the immunogenic peptide.

As another example of lipid priming of CTL responses, *E. coli* lipoproteins, such as tripalmitoyl-S-glycerylcysteinylserine (P<sub>3</sub>CSS) can be used to prime virus specific CTL when covalently attached to an appropriate peptide. (*See, Deres, et al.*,

*Nature* 342:561, 1989). Peptides of the invention can be coupled to P<sub>3</sub>CSS, for example, and the lipopeptide administered to an individual to specifically prime a CTL response to the target antigen. Moreover, because the induction of neutralizing antibodies can also be primed with P<sub>3</sub>CSS-conjugated epitopes, two such compositions can be combined to more effectively elicit both humoral and cell-mediated responses to infection.

In addition, additional amino acids can be added to the termini of a peptide to provide for ease of linking peptides one to another, for coupling to a carrier support, or larger peptide, for modifying the physical or chemical properties of the peptide or oligopeptide, or the like. Amino acids such as tyrosine, cysteine, lysine, glutamic or aspartic acid, or the like, can be introduced at the C- or N-terminus of the peptide or oligopeptide, particularly class I peptides. However, it is to be noted that modification at the carboxyl terminus of a CTL epitope may, in some cases, alter binding characteristics of the peptide. In addition, the peptide or oligopeptide sequences can differ from the natural sequence by being modified by terminal-NH<sub>2</sub> acylation, *e.g.*, by alkanoyl (C<sub>1</sub>-C<sub>20</sub>) or thioglycolyl acetylation, terminal-carboxyl amidation, *e.g.*, ammonia, methylamine, *etc.* In some instances these modifications may provide sites for linking to a support or other molecule.

#### **IV.L. Administration of Vaccines for Therapeutic or Prophylactic Purposes**

The peptides of the present invention and pharmaceutical and vaccine compositions of the invention are useful for administration to mammals, particularly humans, to treat and/or prevent HBV infection. Vaccine compositions containing the peptides of the invention are administered to a patient susceptible to or otherwise at risk for HBV infection to elicit an immune response against HBV antigens and thus enhance the patient's own immune response capabilities. In therapeutic applications, compositions are administered to a patient in an amount sufficient to elicit an effective CTL response to the virus antigen and to cure or at least partially arrest or slow symptoms and/or complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, *e.g.*, the particular composition administered, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician.

The vaccine compositions of the invention may also be used purely as prophylactic agents. Vaccine compositions containing the peptide epitopes of the

invention are administered to a patient susceptible to, or otherwise at risk for, HBV infection to elicit an immune response against HBV antigens and thus enhance the patient's own immune response capabilities following exposure to HBV. Generally the dosage for an initial prophylactic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1000  $\mu\text{g}$  and the higher value is about 10,000; 20,000; 30,000; or 50,000  $\mu\text{g}$ . Dosage values for a human typically range from about 500  $\mu\text{g}$  to about 50,000  $\mu\text{g}$  per 70 kilogram patient. This is followed by boosting dosages of between about 1.0  $\mu\text{g}$  to about 50,000  $\mu\text{g}$  of peptide administered at defined intervals from about four weeks to six months after the initial administration of vaccine.

The immunogenicity of the vaccine may be assessed by measuring the specific activity of CTL and/or HTL obtained from a sample of the patient's blood.

As noted above, peptides comprising CTL or HTL epitopes of the invention induce immune responses when presented by HLA molecules and contacted with a CTL or HTL specific for an epitope comprised by the peptide. The manner in which the peptide is contacted with the CTL or HTL is not critical to the invention. For instance, the peptide can be contacted with the CTL or HTL either *in vivo* or *in vitro*. If the contacting occurs *in vivo*, the peptide itself can be administered to the patient, or other vehicles, *e.g.*, DNA vectors encoding one or more peptides, viral vectors encoding the peptide(s), liposomes and the like, can be used, as described herein.

For pharmaceutical compositions, the immunogenic peptides of the invention, or DNA encoding them, are generally administered to an individual already infected with HBV. The peptides or DNA encoding them can be administered individually or as fusions of one or more peptide sequences. Those in the incubation phase or the acute phase of infection can be treated with the immunogenic peptides separately or in conjunction with other treatments, as appropriate.

For therapeutic use, administration should generally begin at the first diagnosis of HBV infection. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. In chronic infection, loading doses followed by boosting doses may be required.

Treatment of an infected individual with the compositions of the invention may hasten resolution of the infection in acutely infected individuals. For those individuals susceptible (or predisposed) to developing chronic infection, the compositions are particularly useful in methods for preventing the evolution from acute to chronic infection. Where susceptible individuals are identified prior to or during infection, the

composition can be targeted to them, thus minimizing the need for administration to a larger population.

The peptide or other compositions used for the treatment or prophylaxis of HBV infection can be used, *e.g.*, in persons who have not manifested symptoms of disease but who act as a disease vector. In this context, it is generally important to provide an amount of the peptide epitope delivered by a mode of administration sufficient to effectively stimulate a cytotoxic T cell response; compositions which stimulate helper T cell responses can also be given in accordance with this embodiment of the invention.

The dosage for an initial immunization (*i.e.*, therapeutic or prophylactic administration) generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1000  $\mu\text{g}$  and the higher value is about 10,000; 20,000; 30,000; or 50,000  $\mu\text{g}$ . Dosage values for a human typically range from about 500  $\mu\text{g}$  to about 50,000  $\mu\text{g}$  per 70 kilogram patient. Boosting dosages of between about 1.0  $\mu\text{g}$  to about 50000  $\mu\text{g}$  of peptide pursuant to a boosting regimen over weeks to months may be administered depending upon the patient's response and condition as determined by measuring the specific activity of CTL and/or HTL obtained from the patient's blood. The peptides and compositions of the present invention may be employed in serious disease states, that is, life-threatening or potentially life threatening situations. In such cases, as a result of the minimal amounts of extraneous substances and the relative nontoxic nature of the peptides in preferred compositions of the invention, it is possible and may be felt desirable by the treating physician to administer substantial excesses of these peptide compositions relative to these stated dosage amounts.

Thus, for treatment of chronic infection, a representative dose is in the range disclosed above, namely where the lower value is about 1, 5, 50, 500, or 1000  $\mu\text{g}$  and the higher value is about 10,000; 20,000; 30,000; or 50,000  $\mu\text{g}$ , preferably from about 500  $\mu\text{g}$  to about 50,000  $\mu\text{g}$  per 70 kilogram patient. Initial doses followed by boosting doses at established intervals, *e.g.*, from four weeks to six months, may be required, possibly for a prolonged period of time to effectively immunize an individual. In the case of chronic infection, administration should continue until at least clinical symptoms or laboratory tests indicate that the viral infection has been eliminated or substantially abated and for a period thereafter. The dosages, routes of administration, and dose schedules are adjusted in accordance with methodologies known in the art.

The pharmaceutical compositions for therapeutic treatment are intended for parenteral, topical, oral, intrathecal, or local administration. Preferably, the

pharmaceutical compositions are administered parentally, *e.g.*, intravenously, subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic peptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, *e.g.*, water, buffered water, 0.8% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservatives, and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, *etc.*

The concentration of peptides of the invention in the pharmaceutical formulations can vary widely, *i.e.*, from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, *etc.*, in accordance with the particular mode of administration selected.

A human unit dose form of the peptide composition is typically included in a pharmaceutical composition that comprises a human unit dose of an acceptable carrier, preferably an aqueous carrier, and is administered in a volume of fluid that is known by those of skill in the art to be used for administration of such compositions to humans (*see, e.g., Remington's Pharmaceutical Sciences*, 17<sup>th</sup> Edition, A. Gennaro, Editor, Mack Publishing Co., Easton, Pennsylvania, 1985)

The peptides of the invention may also be administered via liposomes, which serve to target the peptides to a particular tissue, such as lymphoid tissue, or to target selectively to infected cells, as well as to increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations, the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver

the peptide compositions. Liposomes for use in accordance with the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, *e.g.*, liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, *e.g.*, Szoka, *et al.*, *Ann. Rev. Biophys. Bioeng.* 9:467 (1980), and U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

For targeting cells of the immune system, a ligand to be incorporated into the liposome can include, *e.g.*, antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide may be administered intravenously, locally, topically, *etc.* in a dose which varies according to, *inter alia*, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

For aerosol administration, the immunogenic peptides are preferably supplied in finely divided form along with a surfactant and propellant. Typical percentages of peptides are 0.01%-20% by weight, preferably 1%-10%. The surfactant must, of course, be nontoxic, and preferably soluble in the propellant. Representative of such agents are the esters or partial esters of fatty acids containing from 6 to 22 carbon atoms, such as caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed esters, such as mixed or natural glycerides may be employed. The surfactant may constitute 0.1%-20% by weight of the composition, preferably 0.25-5%. The balance of the composition is ordinarily propellant. A carrier can also be included, as desired, as with, *e.g.*, lecithin for intranasal delivery.

#### IV.M. Kits

The peptide and nucleic acid compositions of this invention can be provided in kit form together with instructions for vaccine administration. Typically the kit would include desired peptide compositions in a container, preferably in unit dosage form and instructions for administration. An alternative kit would include a minigene construct with desired nucleic acids of the invention in a container, preferably in unit dosage form together with instruction for administration. Lymphokines such as IL-2 or IL-12 may also be included in the kit. Other kit components that may also be desirable include, for example, a sterile syringe, booster dosages, and other desired excipients.

The invention will be described in greater detail by way of specific examples. The following examples are offered for illustrative purposes, and are not intended to limit the invention in any manner. Those of skill in the art will readily recognize a variety of non-critical parameters that can be changed or modified to yield alternative embodiments in accordance with the invention.

#### V. EXAMPLES

##### Example 1: HLA Class I and Class II Binding Assays

The following example of peptide binding to HLA molecules demonstrates quantification of binding affinities of HLA class I and class II peptides. Binding assays can be performed with peptides that are either motif-bearing or not motif-bearing.

Epstein-Barr virus (EBV)-transformed homozygous cell lines, fibroblasts, CIR, or 721.22 transfectants were used as sources of HLA class I molecules. These cells were maintained *in vitro* by culture in RPMI 1640 medium supplemented with 2mM L-glutamine (GIBCO, Grand Island, NY), 50μM 2-ME, 100μg/ml of streptomycin, 100U/ml of penicillin (Irvine Scientific) and 10% heat-inactivated FCS (Irvine Scientific, Santa Ana, CA). Cells were grown in 225-cm<sup>2</sup> tissue culture flasks or, for large-scale cultures, in roller bottle apparatuses. The specific cell lines routinely used for purification of MHC class I and class II molecules are listed in Table XXIV.

Cell lysates were prepared and HLA molecules purified in accordance with disclosed protocols (Sidney *et al.*, *Current Protocols in Immunology* 18.3.1 (1998); Sidney, *et al.*, *J. Immunol.* 154:247 (1995); Sette, *et al.*, *Mol. Immunol.* 31:813 (1994)). Briefly, cells were lysed at a concentration of 10<sup>8</sup> cells/ml in 50 mM Tris-HCl, pH 8.5,

containing 1% Nonidet P-40 (Fluka Biochemika, Buchs, Switzerland), 150 mM NaCl, 5 mM EDTA, and 2 mM PMSF. Lysates were cleared of debris and nuclei by centrifugation at 15,000 x g for 30min.

HLA molecules were purified from lysates by affinity chromatography. Lysates prepared as above were passed twice through two pre-columns of inactivated Sepharose CL4-B and protein A-Sepharose. Next, the lysate was passed over a column of Sepharose CL-4B beads coupled to an appropriate antibody. The antibodies used for the extraction of HLA from cell lysates are listed in Table XXV. The anti-HLA column was then washed with 10-column volumes of 10mM Tris-HCL, pH 8.0, in 1% NP-40, PBS, 2-column volumes of PBS, and 2-column volumes of PBS containing 0.4% n-octylglucoside. Finally, MHC molecules were eluted with 50mM diethylamine in 0.15M NaCl containing 0.4% n-octylglucoside, pH 11.5. A 1/25 volume of 2.0M Tris, pH 6.8, was added to the eluate to reduce the pH to ~8.0. Eluates were then be concentrated by centrifugation in Centriprep 30 concentrators at 2000 rpm (Amicon, Beverly, MA). Protein content was evaluated by a BCA protein assay (Pierce Chemical Co., Rockford, IL) and confirmed by SDS-PAGE.

A detailed description of the protocol utilized to measure the binding of peptides to Class I and Class II MHC has been published (Sette *et al.*, *Mol. Immunol.* 31:813, 1994; Sidney *et al.*, in *Current Protocols in Immunology*, Margulies, Ed., John Wiley & Sons, New York, Section 18.3, 1998). Briefly, purified MHC molecules (5 to 500nM) were incubated with various unlabeled peptide inhibitors and 1-10nM <sup>125</sup>I-radiolabeled probe peptides for 48h in PBS containing 0.05% Nonidet P-40 (NP40) (or 20% w/v digitonin for H-2 IA assays) in the presence of a protease inhibitor cocktail. The final concentrations of protease inhibitors (each from CalBioChem, La Jolla, CA) were 1 mM PMSF, 1.3 nM 1.10 phenanthroline, 73 μM pepstatin A, 8mM EDTA, 6mM N-ethylmaleimide (for Class II assays), and 200 μM N alpha-p-tosyl-L-lysine chloromethyl ketone (TLCK). All assays were performed at pH 7.0 with the exception of DRB1\*0301, which was performed at pH 4.5, and DRB1\*1601 (DR2w21 ·1) and DRB4\*0101 (DRw53), which were performed at pH 5.0. pH was adjusted as described elsewhere (see Sidney *et al.*, in *Current Protocols in Immunology*, Margulies, Ed., John Wiley & Sons, New York, Section 18.3, 1998).

Following incubation, MHC-peptide complexes were separated from free peptide by gel filtration on 7.8 mm x 15 cm TSK200 columns (TosoHaas 16215,



Montgomeryville, PA), eluted at 1.2 mls/min with PBS pH 6.5 containing 0.5% NP40 and 0.1%  $\text{NaN}_3$ . Because the large size of the radiolabeled peptide used for the DRB1\*1501 (DR2w2.1) assay makes separation of bound from unbound peaks more difficult under these conditions, all DRB1\*1501 (DR2w2.1) assays were performed using a 7.8mm x 30cm TSK2000 column eluted at 0.6 mls/min. The eluate from the TSK columns was passed through a Beckman 170 radioisotope detector, and radioactivity was plotted and integrated using a Hewlett-Packard 3396A integrator, and the fraction of peptide bound was determined.

Radiolabeled peptides were iodinated using the chloramine-T method. The specific radiolabeled probe peptide utilized in each assay, and its assay specific  $\text{IC}_{50}$  nM, is summarized in Table XXIV. Typically, in preliminary experiments, each MHC preparation was titrated in the presence of fixed amounts of radiolabeled peptides to determine the concentration of HLA molecules necessary to bind 10-20% of the total radioactivity. All subsequent inhibition and direct binding assays were performed using these HLA concentrations.

Since under these conditions  $[\text{label}] < [\text{HLA}]$  and  $\text{IC}_{50} \geq [\text{HLA}]$ , the measured  $\text{IC}_{50}$  values are reasonable approximations of the true  $K_D$  values. Peptide inhibitors are typically tested at concentrations ranging from 120  $\mu\text{g/ml}$  to 1.2  $\text{ng/ml}$ , and are tested in two to four completely independent experiments. To allow comparison of the data obtained in different experiments, a relative binding figure is calculated for each peptide by dividing the  $\text{IC}_{50}$  of a positive control for inhibition by the  $\text{IC}_{50}$  for each tested peptide (typically unlabeled versions of the radiolabeled probe peptide). For database purposes, and inter-experiment comparisons, relative binding values are compiled. These values can subsequently be converted back into  $\text{IC}_{50}$  nM values by dividing the  $\text{IC}_{50}$  nM of the positive controls for inhibition by the relative binding of the peptide of interest. This method of data compilation has proven to be the most accurate and consistent for comparing peptides that have been tested on different days, or with different lots of purified MHC.

Because the antibody used for HLA-DR purification (LB3.1) is  $\alpha$ -chain specific,  $\beta_1$  molecules are not separated from  $\beta_3$  (and/or  $\beta_4$  and  $\beta_5$ ) molecules. The  $\beta_1$  specificity of the binding assay is obvious in the cases of DRB1\*0101 (DR1), DRB1\*0802 (DR8w2), and DRB1\*0803 (DR8w3), where no  $\beta_3$  is expressed. It has also been demonstrated for DRB1\*0301 (DR3) and DRB3\*0101 (DR52a), DRB1\*0401 (DR4w4), DRB1\*0404

(DR4w14), DRB1\*0405 (DR4w15), DRB1\*1101 (DR5), DRB1\*1201 (DR5w12), DRB1\*1302 (DR6w19) and DRB1\*0701 (DR7). The problem of  $\beta$  chain specificity for DRB1\*1501 (DR2w2.1), DRB5\*0101 (DR2w2.2), DRB1\*1601 (DR2w21.1), DRB5\*0201 (DR2w21.3), and DRB4\*0101 (DRw53) assays is circumvented by the use of fibroblasts. Development and validation of assays with regard to DR $\beta$  molecule specificity have been described previously (see, *e.g.*, Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998).

Binding assays as outlined above may be used to analyze supermotif and/or motif-bearing epitopes as, for example, described in Example 2.

#### Example 2. Identification of Conserved HLA Supermotif CTL Candidate Epitopes

Vaccine compositions of the invention may include multiple epitopes that comprise multiple HLA supermotifs or motifs to achieve broad population coverage. This example illustrates the identification of supermotif-bearing epitopes for the inclusion in such a vaccine composition. Calculation of population coverage was performed using the strategy described below. Epitopes were then selected to bear an HLA-A2, -A3, or -B7 supermotif or an HLA-A1 or -A24 motif.

#### *Computer searches and algorithms for identification of supermotif and/or motif-bearing epitopes*

Computer searches for epitopes bearing HLA Class I or Class II supermotifs or motifs were performed as follows. All translated HBV isolate sequences were analyzed using a text string search software program, *e.g.*, MotifSearch 1.4 (D. Brown, San Diego) to identify potential peptide sequences containing appropriate HLA binding motifs; alternative programs are readily produced in accordance with information in the art in view of the motif/supermotif disclosure herein. Furthermore, such calculations can be made mentally. Identified A2-, A3-, and DR-supermotif sequences were scored using polynomial algorithms to predict their capacity to bind to specific HLA-Class I or Class II molecules. These polynomial algorithms take into account both extended and refined motifs (that is, to account for the impact of different amino acids at different positions), and are essentially based on the premise that the overall affinity (or  $\Delta G$ ) of peptide-HLA molecule interactions can be approximated as a linear polynomial function of the type:

$$“\Delta G” = a_{1i} \times a_{2i} \times a_{3i} \dots \times a_{ni}$$

where  $a_{ji}$  is a coefficient which represents the effect of the presence of a given amino acid ( $j$ ) at a given position ( $i$ ) along the sequence of a peptide of  $n$  amino acids. The crucial assumption of this method is that the effects at each position are essentially independent of each other (i.e., independent binding of individual side-chains). When residue  $j$  occurs at position  $i$  in the peptide, it is assumed to contribute a constant amount  $j_i$  to the free energy of binding of the peptide irrespective of the sequence of the rest of the peptide. This assumption is justified by studies from our laboratories that demonstrated that peptides are bound to MHC and recognized by T cells in essentially an extended conformation (data omitted herein).

The method of derivation of specific algorithm coefficients has been described in Gulukota *et al.*, *J. Mol. Biol.* 267:1258-126, 1997; (see also Sidney *et al.*, *Human Immunol.* 45:79-93, 1996; and Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). Briefly, for all  $i$  positions, anchor and non-anchor alike, the geometric mean of the average relative binding (ARB) of all peptides carrying  $j$  is calculated relative to the remainder of the group, and used as the estimate of  $j_i$ . For Class II peptides, if multiple alignments are possible, only the highest scoring alignment is utilized, following an iterative procedure. To calculate an algorithm score of a given peptide in a test set, the ARB values corresponding to the sequence of the peptide are multiplied. If this product exceeds a chosen threshold, the peptide is predicted to bind. Appropriate thresholds are chosen as a function of the degree of stringency of prediction desired.

#### *Selection of HLA-A2 supertype cross-reactive peptides*

Complete sequences from 20 HBV isolates were aligned, then scanned, utilizing a customized computer program, to identify conserved 9- and 10-mer sequences containing the HLA-A2-supertype main anchor specificity.

A total of 150 conserved and motif-positive sequences were identified. These peptides were then evaluated for the presence of A\*0201 preferred secondary anchor residues using an A\*0201-specific polynomial algorithm. A total of 85 conserved, motif-positive sequences were selected and synthesized.

These 85 conserved, motif-containing 9- and 10-mer peptides were then tested for their capacity to bind purified HLA-A\*0201 molecules in vitro. Thirty-four peptides were found to be good A\*0201 binders ( $IC_{50} \leq 500$  nM); 15 were high binders ( $IC_{50} \leq 50$  nM) and 19 were intermediate binders ( $IC_{50}$  of 50-500 nM) (Table XXVI).

In the course of independent analyses, 25 conserved, HBV-derived, 8 or 11-mer sequences with appropriate A2-supertype main anchors were also synthesized and tested for A\*0201 binding. One peptide, HBV env 259 11-mer (peptide 1147.14), bound A\*0201 with an IC<sub>50</sub> of 500 nM, or less, and has been included in Table XXVI. Also shown in Table XXVI is an analog peptide, representing a single substitution of the HBV pol 538 9-mer peptide, which binds A\*0201 with an IC<sub>50</sub> of 5.1 nM (see below).

Thirty of the 36 A\*0201 binders were subsequently tested for the capacity to bind to additional A2-supertype alleles (A\*0202, A\*0203, A\*0206, and A\*6802). As shown in Table XXVI, 15/30 (50%) peptides were found to be A2-supertype cross-reactive binders, binding at least 3 of the 5 A2-supertype alleles tested. These 15 peptides were selected for further analysis.

#### *Selection of HLA-A3 supermotif-bearing epitopes*

The sequences from the same 20 isolates were also examined for the presence of conserved peptides with the HLA-A3-supermotif primary anchors. A total of 80 conserved 9- or 10-mer motif-containing sequences were identified. Further analysis using the A03 and A11 algorithms identified 40 sequences which scored high in either or both algorithms. Thirty-six of the corresponding peptides were synthesized and tested for binding to HLA-A\*03 and HLA-A\*11, the two most prevalent A3-supertype alleles.

Twenty-three peptides were identified which bound A3 and/or A11 with affinities or IC<sub>50</sub> values of ≤ 500 nM (Table XXVII).

In the course of an independent series of studies 30 HBV-derived 8-mer, and 24 11-mer sequences, conserved in 75% or more of the isolates, were selected and tested for A\*03 and A\*11 binding. Four 8-mers and 9 11-mers were found to bind either or both alleles (Table XXVII). Finally, four 9-mer, and one 10-mer, conserved HBV-derived peptides not selected using the search criteria outlined above, but which have been shown to bind A\*03 and/or A\*11, have been identified, and are included in Table XXVII. Two of these peptides contain non-canonical anchors (peptides 20.0131, and 20.0130), and the other 3 are algorithm negative (peptides 1142.05, 1099.03, and 1090.15).

Thirty-eight of the 41 peptides binding A\*03 and/or A\*11 were subsequently tested for binding crossreactivity to the other common A3-supertype alleles (A\*3101, A\*3301, and A\*6801). It was found that 17 of these peptides were A3-supertype cross-reactive, binding at least 3 of the 5 A3-supertype alleles tested (Table XXVII).

### *Selection of HLA-B7 supermotif bearing epitopes*

When the same 20 isolates were also analyzed for the presence of conserved 9- or 10-mer peptides with the HLA-B7-supermotif, 46 sequences were identified. Thirty-four of the corresponding peptides were synthesized and tested for binding to HLA-B\*0702, the most common B7-supertype allele. Nine peptides bound B\*0702 with an  $IC_{50}$  value of  $\leq 500$  nM (Table XXVIII). These 9 peptides were then tested for binding to other common B7-supertype alleles (B\*3501, B\*51, B\*5301, and B\*5401). Five of the 9 B\*0702 binders were capable of binding to 3 or more of the 5 B7-supertype alleles tested.

In separate studies investigating the secondary anchor requirements of B7-supertype alleles, all available peptides with the B7-supermotif were tested for binding to all B7 supertype alleles. As a result, all 34 peptides described above were also tested for binding to other B7-supertype alleles. These experiments identified an additional 10 peptides which bound at least one B7-supertype allele with an  $IC_{50}$  value  $\leq 500$  nM, including 2 peptides which bound 3 or more alleles. These 10 peptides are also shown in Table XXVIII.

Because of the low numbers of conserved B7-supertype degenerate HBV-derived 9- and 10-mer peptides, compared to the A2- and A3-supertypes, the 20 isolates were again examined to identify conserved, motif-containing 8- and 11-mers. This re-scan identified 51 peptides. Thirty-one of these were synthesized and tested for binding to each of the 5 most common B7-supertype alleles. Nineteen 8- and 11-mer peptides bound with high or intermediate affinity to at least one B7-supertype allele ( $IC_{50} \leq 500$  nM) (Table XXVIII). Two peptides were degenerate binders, binding 3 of the 5 alleles tested.

In summary, a total of 9 HBV-derived peptides, conserved in 75% or more of the isolates analyzed, have been identified which are degenerate B7-supertype binders (Table XXVIII).

### *Selection of A1 and A24 motif-bearing epitopes*

To further increase population coverage, HLA-A1 and -A24 epitopes have been incorporated into the present analysis. A1 is, on average, present in 12%, and A24 is present in approximately 29%, of the population across five different major ethnic groups (Caucasian, North American Black, Chinese, Japanese, and Hispanic). Combined, these

alleles would be represented with an average frequency of 39% in these same populations. The total coverage across the major ethnicities when A1 and A24 are combined with the coverage of the A2-, A3- and B7-supertype alleles is >95.4%; by comparison, coverage by combining the A2-, A3-, and B7-supertypes is 86.2%.

5        Systematic analyses of HBV for A1 and A24 binders have yet to be completed . However, in the course of independent studies, 15 conserved HBV-derived peptides have been identified that bind A\*0101 with IC<sub>50</sub> less than 500 nM (Table XXIX); 7 of these bind with IC<sub>50</sub> less than 100 nM . In a similar context, 14 conserved A\*2402 binding HBV-derived peptides have also been identified, 6 of which bind A\*2402 with IC<sub>50</sub> less  
10    than 100 nM (Table XXIX).

### Example 3: Confirmation of Immunogenicity

#### *Evaluation of A\*0201 immunogenicity*

15        The immunogenicity analysis of the 15 HBV-derived HLA-A2 supertype cross-reactive peptides identified above is summarized in Table XXX. Peptides were screened for immunogenicity in at least one of three systems. Peptides were screened for the induction of primary antigen-specific CTL *in vitro* using human PBMC (Wentworth *et al.*, *Molec. Immunol.* 32:603, 1995); this data is indicated as "primary CTL" in Table XXX.

20        The protocol for *in vitro* induction of primary antigen-specific CTL from human PBMC has been described by Wentworth *et al.* (Wentworth *et al.*, *Molec. Immunol.* 32:603, 1995). PBMC from normal donors which had been enriched for CD8<sup>+</sup> T cells were incubated with peptide loaded antigen-presenting cells (SAC-I activated PBMCs) in the presence of IL-7. After seven days cultures were restimulated using irradiated  
25    autologous adherent cells pulsed with peptide and then tested for cytotoxic activity seven days later.

      In addition, HLA transgenic mice were used to evaluate peptide immunogenicity; this data is indicated as "transgenic CTL" in Table XXX. Previous studies have shown that CTL induced in A\*0201/Kb transgenic mice exhibit specificity similar to CTL  
30    induced in humans (Vitiello *et al.*, *J. Exp. Med.* 173:1007, 1991; Wentworth *et al.*, *Eur. J. Immunol.* 26:97, 1996).

      CTL induction in transgenic mice following peptide immunization has been described by Vitiello *et al.* (Vitiello *et al.*, *J. Exp. Med.* 173:1007, 1991) and Alexander *et al.* (Alexander *et al.*, *J. Immunol.* 159:4753, 1997). Briefly, synthetic peptides (50

µg/mouse) and the helper epitope HBV core 128 (140 µg/mouse) were emulsified in incomplete Freund's adjuvant (IFA) and injected subcutaneously at the base of the tail. Eleven days after injection, splenocytes were incubated in the presence of peptide-loaded syngenic LPS blasts. After six days cultures were assayed for cytotoxic activity using peptide-pulsed targets.

Peptides were also tested for the ability to stimulate recall CTL responses in acutely infected HBV patients (Bertoni *et al.*, *J. Clin. Invest.* 100:503, 1997; Rehmann *et al.*, *J. Clin. Invest.* 97:1655-1665, 1996; Nayersina *et al.*, *J. Immunol.* 150:4659, 1993); these data are indicated as "patient CTL" in Table XXX. Patient immunogenicity data is particularly informative as it indicates that a peptide is recognized during the course of a natural infection. These data demonstrate that a peptide is processed and presented in human cells that would represent the targets for CTL. Moreover, this data is especially relevant for vaccine design as the induction of CTL responses in patients has been correlated to the resolution of infection.

For the evaluation of recall CTL responses, screening was carried out as described by Bertoni *et al.* (Bertoni *et al.*, *J. Clin. Invest.* 100:503, 1997). Briefly, PBMC from patients acutely infected with HBV were cultured in the presence of 10µg/ml of synthetic peptide. After seven days, the cultures were restimulated with peptide. The cultures were assayed for cytotoxic activity on day 14 using target cells pulsed with peptide.

Of the 15 A2 supertype binding peptides, 11 were found to be immunogenic in at least one of the systems utilized. Five of the 11 peptides had previously been identified in the patients with acute HBV (Bertoni *et al.*, *J. Clin. Invest.* 100:503, 1997). Five additional degenerate peptides (1069.06, 1090.77, 1147.14, 927.42 and 927.46) induced CTL responses in HLA-A\*0201 transgenic mice. The 11 immunogenic supertype cross-reactive peptides are encoded by three HBV antigens; core, envelope and polymerase.

This set of 11 immunogenic A2-supermotif-bearing epitopes includes one analog peptide, 1090.77. The wild type peptide, 1090.14, from which this analog is derived is A2-supertype non-cross-reactive, but has been shown to be recognized in recall CTL responses from acute HBV patients, and to be immunogenic in HLA-A\*0201 transgenic mice as well as primary human cultures (Table XXX). Further studies addressing the cross recognition of the wild type peptide 1090.14 and the 1090.77 analog are described in detail below.

In the course of independent analyses, 14 of the non-cross-reactive peptides shown in Table XXXb, including 1090.14, were found to be immunogenic in at least one

system. Five peptides of these peptides were recognized in patients; 4 peptides induced CTL in transgenic mice.

In conclusion, 11 A2-supertype cross-reactive peptides have been identified that are capable of exhibiting immunogenicity in at least one of the three systems examined.

5

#### *Evaluation of A\*03/A11 immunogenicity*

Seven of the 17 A3-supertype cross-reactive peptides have been evaluated for immunogenicity (Table XXXI). As described in the previous section, A3-supermotif-bearing peptides were screened using primary cultures, patient responses, or HLA-A11 transgenic mice (Alexander *et al.*, *J. Immunol.* 159:4753, 1997). With the exception of peptide 1.0219, all of the conserved cross-reactive peptides listed in Table insert table XXXI were found to be immunogenic.

Additionally, a poorly conserved peptide (1150.51; 40% conserved) which exhibits cross-reactive supertype binding was found to be immunogenic in transgenic mice, and has been included in Table XXXI. Two other conserved, but non-cross-reactive, peptides have also been shown to be recognized in acutely infected HBV patients (Bertoni *et al.*, *J. Clin. Invest.* 100:503, 1997). These epitopes are shown in Table XXXI.

It is notable that for 7 of the 8 conserved immunogenic HBV-derived A3-supermotif-bearing epitopes, including all 6 of the cross-reactive peptides, positive data was obtained in patients. These epitopes are predominantly derived from the polymerase protein sequence, with only one epitope being derived from the core protein sequence. While a number of cross-reactive peptides have been identified in the X antigen (Table XXXI), to date these peptides have not been screened for immunogenicity.

In summary, 7 A3-supermotif-bearing, cross-reactive peptides have been identified that are recognized by CTL in acutely infected patients, or induce CTL in HLA-transgenic mice.

#### *Evaluation of B7 immunogenicity*

The immunogenicity studies involving the HBV-derived HLA-B7-supermotif-bearing, cross-reactive peptides is summarized in Table XXXII. HLA-B7 peptides were screened exclusively in human systems measuring responses in either primary cultures or acutely infected HBV patients. Of the 7 degenerate peptides screened, 4 were shown to be immunogenic. One non-crossreactive peptide (XRN<3), 1147.04, was also shown to



be recognized in acutely infected HBV patients (Bertoni *et al.*, *J. Clin. Invest.* 100:503, 1997; see TableXXXII).

In summary, 5 conserved HBV-derived B7-supermotif-bearing epitopes that are recognized in acutely infected HBV patients have been identified. These epitopes afford coverage of 4 different HBV antigens (core, envelope, polymerase and X).

#### Example 4: Implementation of the Extended Supermotif to Improve the Binding Capacity of Native Peptides by Creating Analogs

HLA motifs and supermotifs (comprising primary and/or secondary residues) are useful in preparing highly cross-reactive native peptides, as demonstrated herein. Moreover, the definition of HLA motifs and supermotifs also allows one to engineer highly cross-reactive epitopes by identifying residues within a native peptide sequence which can be analoged, or "fixed", to confer upon a peptide certain characteristics, *e.g.*, greater cross-reactivity within the group of HLA molecules that make-up the supertype, and/or greater binding affinity for some or all of those HLA molecules. Examples of analog peptides that exhibit modulated binding affinity are provided.

#### *Analoging at Primary Anchor Residues*

It has been shown that class I peptide ligands can be modified, or "fixed" to increase their binding affinity and/or degeneracy (Sidney *et al.*, *J. Immunol.* 157:3480, 1996). These fixed peptides may also demonstrate increased immunogenicity and crossreactive recognition by T cells specific for the wild type epitope (Parkhurst *et al.*, *J. Immunol.* 157:2539, 1996; Pogue *et al.*, *Proc. Natl. Acad. Sci. USA* 92:8166, 1995). Specifically, the main anchors of A2 supertype peptides may be "fixed", or analoged, to L or V (or M, if natural) at position 2, and V at the C-terminus. As indicated in Table XXVI, 9 of the 14 A2-supertype cross-reactive binding peptides are "fixable" by these criteria, as are 16 of the 21 non-cross-reactive binders. Ideal candidates for fixing would be peptides which bind at least 3 A2-supertype allele-specific molecules with  $IC_{50} \leq 5000$  nM.

An example of the efficacy of this strategy to generate more broadly cross-reactive epitopes is provided by the case of peptide 1090.14 (Table XXVI). Previously, this peptide was shown to be highly immunogenic in each of the systems examined. However, it only exhibits binding to a single A2-supertype allele-specific molecule,

A\*0201. The non-crossreactive binding capacity of this epitope limits the population coverage and consequently the value of including this peptide in a candidate vaccine. In an effort to increase binding affinity and cross-reactivity the C-terminus of peptide 1090.14 was altered from 'alanine' to the A2-supermotif preferred residue 'valine'. This change resulted in a dramatic (40-fold) increase in binding capacity for A\*0201 (from 200 nM to 5.1 nM), but also produced a peptide capable of binding 3 other A2-supertype allele-specific molecules. (see peptide 1090.77, Table XXVI).

Studies with HLA-A\*0201 transgenic mice have shown that the CTL response from mice immunized with the 1090.77 peptide recognize target cells loaded with either the naturally occurring peptide 1090.14 or the valine-substituted analog (*i.e.*, 1090.77). In fact, the lysis effected by 1090.77 induced CTL was indistinguishable regardless whether the analog or the wild-type sequence was used to load the target cells (B. Livingston, unpublished data).

The relevance of these observations for the design of vaccine constructs is indicated by studies in which chronic HBV patients were treated with the potent viral replication inhibitor, lamivudine. Extended therapy with lamivudine resulted in the selection of drug-resistant strains of HBV that have a substitution of valine for methione at position 2 in the 1090.14 epitope, suggesting that epitope-based vaccines used in combination with lamivudine may need to have the ability to induce CTL responses that recognize both wild type and mutant sequences.

To demonstrate that cross-recognition is possible between the native peptide (1090.14), the analog peptide, and the lamivudine induced mutant M2 peptide, CTL were generated using the 1090.77 analog peptide. These CTL cultures were then stimulated with either the wild type peptide (1090.14), or the lamivudine induced mutant M2 peptide. The ability of these CTL to then lyse target cells loaded with either the wild type, or the lamivudine induced mutant peptide was then assayed. Target cells presenting either peptide were similarly lysed by either CTL culture (Table XXVI).

These studies demonstrate how analoging a peptide can result in dramatically increased HLA-A2 supertype degeneracy while still allowing cross-recognition between wildtype and mutant epitopes. More specifically, these results indicate that a vaccine utilizing the analog peptide 1090.77 should stimulate a response that will recognize both wild-type and lamivudine-resistant strains of HBV.

Similarly, analogs of HLA-A3 supermotif-bearing epitopes may also be generated. For example, peptides could be analogued to possess a preferred V at position

2, and R or K at the C-terminus. Twelve of the A3-supertype degenerate peptides identified in Table XXVII are candidates for main anchor fixing, as are 19 of the 24 non-cross-reactive binders.

5 Analog peptides are initially tested for binding to A\*03 and A\*11, and those that demonstrate equivalent, or improved, binding capacity relative to the parent peptide would then be tested for A3-supertype cross-reactivity. Analogs demonstrating improved cross-reactivity are then further evaluated for immunogenicity, as necessary.

10 Typically, it is more difficult to identify B7 supermotif-bearing epitopes. As in the cases of A2- and A3-supertype epitopes, a peptide analoging strategy can be utilized to generate additional B7 supermotif-bearing epitopes with increased cross-reactive binding. In general, B7 supermotif-bearing peptides should be fixed to possess P in position 2, and I at their C-terminus.

15 Analogs representing primary anchor single amino acid residues substituted with I residues at the C-terminus of two different B7-like peptides (HBV env 313 and HBV pol 541) were synthesized and tested for their B7-supertype binding capacity. It was found that the I substitution had an overall positive effect on binding affinity and/or cross-reactivity in both cases. In the case of HBV env 313 the I9 (I at C-terminal position 9) replacement was effective in increasing cross-reactivity from 4 to 5 alleles bound by virtue of an almost 400-fold increase B\*5401 binding affinity. In the case of HBV pol 20 541, increased cross-reactivity was similarly achieved by a substantial increase in B\*5401 binding. Also, significant gains in binding affinity for B\*0702, B51, and B\*5301 were observed with the HBV pol 541 I9 analog.

#### *Analoging at Secondary Anchor Residues*

25 Moreover, HLA supermotifs are of value in engineering highly cross-reactive peptides by identifying particular residues at secondary anchor positions that are associated with such cross-reactive properties. Demonstrating this, the capacity of a second set of peptides representing discreet single amino acid substitutions at positions one and three of five different B7-supertype binding peptides were synthesized and tested 30 for their B-7 supertype binding capacity. In 4/4 cases the effect of replacing the native residue at position 1 with the aromatic residue F (an "F1" substitution) resulted in an increase in cross-reactivity, compared to the parent peptide, and, in most instances, binding affinity was increased three-fold or better (Table XXVIII). More specifically, for HBV env 313, MAGE2 170, and HBV core 168 complete supertype cross-reactivity was

achieved with the F1 substitution analogs. These gains were achieved by dramatically increasing B\*5401 binding affinity. Also, gains in affinity were noted for other alleles in the cases of HBV core 168 (B\*3501 and B\*5301) and MAGE2 170 (B\*3501, B51 and B\*5301). Finally, in the case of MAGE3 196, the F1 replacement was effective in increasing cross-reactivity because of gains in B\*0702 binding. An almost 70-fold increase in B51 binding capacity was also noted.

Two analogs were also made using the supermotif positive F substitution at position three (an “F3” substitution). In both instances increases in binding affinity and cross-reactivity were achieved. Specifically, in the case of HBV pol 541, the F3 substitution was effective in increasing cross-reactivity by virtue of its effect on B\*5401 binding. In the case of MAGE3 196, complete supertype cross-reactivity was achieved by increasing B\*0702 and B\*3501 binding capacity. Also, in the case of MAGE3 196, it is notable that increases in binding capacity between 40- and 5000-fold were obtained for B\*3501, B51, B\*5301, and B\*5401.

In conclusion, these data demonstrate that by the use of even single amino acid substitutions, it is possible to increase the binding affinity and/or cross-reactivity of peptide ligands for HLA supertype molecules.

#### Example 5: Identification of conserved HBV-derived sequences with HLA-DR binding motifs

Peptide epitopes bearing an HLA class II supermotif or motif may also be identified as outlined below using methodology similar to that described in Examples 1-3.

#### *Selection of HLA-DR-supermotif-bearing epitopes*

HLA-Class II molecules bind peptides typically between 12 and 20 residues in length. However, similar to HLA-Class I, the specificity and energy of interaction is usually contained within a short core region of about 9 residues. Most DR molecules share an overlapping specificity within this 9-mer core in which a hydrophobic residue in position 1 (P1) is the main anchor (O’Sullivan *et al.*, *J. Immunol.* 147:2663, 1991; Southwood *et al.*, *J. Immunol.* 160:3363, 1998). The presence of small or hydrophobic residues in position 6 (P6) is also important for most DR-peptide interactions. This overlapping P1-P6 specificity, within a 9-mer core region, has been defined as the DR-supermotif. Unlike Class I molecules, DR molecules are open at both ends of the binding groove, and can therefore accommodate longer peptides of varying length. Indeed, while

most of the energy of peptide-DR interactions appears to be contributed by the core region, flanking residues appear to be important for high affinity interactions. Also, although not strictly necessary for MHC binding, flanking residues are clearly necessary in most instances for T cell recognition.

To identify HBV-derived DR cross-reactive HTL epitopes, the same 20 HBV polyproteins that were scanned for the identification of HLA Class I motif sequences were scanned for the presence of sequences with motifs for binding HLA-DR. Specifically, 15-mer sequences comprised of a DR-supermotif containing 9-mer core, and three residue N- and C-terminal flanking regions, were selected. It was also required that 100% of the 15-mer sequence be conserved in at least 85% (17/20) of the HBV strains scanned. Using these criteria, 36 non-redundant sequences were identified. Thirty-five of these peptides were subsequently synthesized.

Algorithms for predicting peptide binding to DR molecules have also been developed (Southwood *et al.*, *J. Immunol.* 160:3363, 1998). These algorithms, specific for individual DR molecules, allow the scoring and ranking of 9-mer core regions. Using selection tables, it has been found that these algorithms efficiently select peptide sequences with a high probability of binding the appropriate DR molecule. Additionally, it has been found that running algorithms, specifically those for DR1, DR4w4, and DR7, sequentially can efficiently select DR cross-reactive peptides.

To see if these algorithms would identify additional peptides, the same HBV polyproteins used above were re-scanned for the presence of 15-mer peptides where 100% of the 9-mer core region was 85% (17/20 strains) conserved. Next, the 9-mer core region of each of these peptides was scored using the DR1, DR4w4, and DR7 algorithms. As a result, 8 additional sequences were identified and synthesized.

In summary, 44 15-mer peptides in which a 9-mer core region contained the DRsupermotif, or was selected using an algorithm predicting DR-binding sequences, were identified. Forty-three of these peptides were synthesized (Table XXXIII).

While performing the analyses of HBV-derived peptides described above, 9 peptides predicted on the basis of their DR1, DR4w4, and DR7 algorithm profiles to be DR-cross-reactive binding peptides, but which have 9-mer core regions that are only 80% conserved, were also identified. An additional peptide which contains a DR-supermotif core region that is 95% conserved, but is located only one residue removed from the N-terminus, was previously synthesized. These 10 peptides were also selected for further analysis, and are shown in Table XXXIII.

Finally, 2 peptides, CF-08 and 1186.25, which are redundant with a peptide selected above (27.0280), were considered for additional analysis. Peptide 1186.25 contains multiple DR-supermotif sequences. Peptide CF-08 is a 20-mer that nests both 27.0280 and 1186.25. These peptides are shown in Table XXXIII.

5 The 55 HBV-derived peptides identified above were tested for their capacity to bind common HLA-DR alleles. To maximize both population coverage, and the relationships between the binding repertoires of most DR alleles (see, *e.g.*, Southwood *et al.*, *J. Immunol.* 160:3363, 1998), peptides were screened for binding to sequential panels of DR assays. The composition of these screening panels, and the phenotypic frequency of associated antigens, are shown in Table XXXIV. All peptides were initially tested for binding to the alleles in the primary panel: DR1, DR4w4, and DR7. Only peptides binding at least 2 of these 3 alleles were then tested for binding in the secondary assays (DR2w2  $\beta$ 1, DR2w2  $\beta$ 2, DR6w19, and DR9). Finally, only peptides binding at least 2 of the 4 secondary panel alleles, and thus 4 of 7 alleles total, were screened for binding in the tertiary assays (DR4w15, DR5w11, and DR8w2).

Upon testing, it was found that 25 of the original 55 peptides (45%) bound two or more of the primary panel alleles. When these 25 peptides were subsequently tested in the secondary assays, 20 were found to bind at least 4 of the 7 DR alleles in the primary and secondary assay panels. Finally, 18 of the 20 peptides passing the secondary screening phase were tested for binding in the tertiary assays. As a result, 12 peptides were shown to bind at least 7 of 10 common HLA-DR alleles. The sequences of these 12 peptides, and their binding capacity for each assay in the primary through tertiary panels, are shown in Table XXXV. Also shown are peptides CF-08 and 857.02, which bound 5/5 and 5/6 of the alleles tested to date, respectively.

25 In summary, 14 peptides, derived from 12 independent regions of the HBV genome, have been identified that are capable of binding multiple HLA-DR alleles. This set of peptides includes at least 2 epitopes each from the Core (Nuc), Pol, and Env antigens.

### 30 *Selection of conserved DR3 motif peptides*

Because HLA-DR3 is an allele that is prevalent in Caucasian, Black, and Hispanic populations, DR3 binding capacity is an important criterion in the selection of HTL epitopes. However, data generated previously indicated that DR3 only rarely cross-reacts with other DR alleles (Sidney *et al.*, *J. Immunol.* 149:2634-2640, 1992; Geluk *et al.*, *J.*

*Immunol.* 152:5742-5748, 1994; Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998).

This is not entirely surprising in that the DR3 peptide-binding motif appears to be distinct from the specificity of most other DR alleles.

To efficiently identify peptides that bind DR3, target proteins were analyzed for conserved sequences carrying one of the two DR3 specific binding motifs reported by Geluk *et al.* (*J. Immunol.* 152:5742-5748, 1994). Eighteen sequences were identified. Eight of these sequences were largely redundant with peptides shown in Table XXXVI, and 3 with peptides that had previously been synthesized for other studies. The 7 unique sequences were synthesized.

Seventeen of the eighteen peptides containing a DR3 motif have been tested for their DR3 binding capacity. Four peptides were found to bind DR3 with an affinity of 1000 nM or better (Table XXXVI).

#### Example 6. Calculation of phenotypic frequencies of HLA-supertypes in various ethnic backgrounds to determine breadth of population coverage

This example illustrates the assessment of the breadth of population coverage of a vaccine composition comprised of multiple epitopes comprising multiple supermotifs and/or motifs.

In order to analyze population coverage, gene frequencies of HLA alleles were determined. Gene frequencies for each HLA allele were calculated from antigen or allele frequencies utilizing the binomial distribution formulae  $gf=1-(\text{SQRT}(1-af))$  (see, *e.g.*, Sidney *et al.*, *Human Immunol.* 45:79-93, 1996). To obtain overall phenotypic frequencies, cumulative gene frequencies were calculated, and the cumulative antigen frequencies derived by the use of the inverse formula  $[af=1-(1-Cgf)^2]$ .

Where frequency data was not available at the level of DNA typing, correspondence to the serologically defined antigen frequencies was assumed. To obtain total potential supertype population coverage no linkage disequilibrium was assumed, and only alleles confirmed to belong to each of the superotypes were included (minimal estimates). Estimates of total potential coverage achieved by inter-loci combinations were made by adding to the A coverage the proportion of the non-A covered population that could be expected to be covered by the B alleles considered (*e.g.*,  $\text{total}=A+B*(1-A)$ ). Confirmed members of the A3-like supertype are A3, A11, A31, A\*3301, and A\*6801. Although the A3-like supertype may potentially include A34, A66, and A\*7401, these alleles were not included in overall frequency calculations. Likewise, confirmed

members of the A2-like supertype family are A\*0201, A\*0202, A\*0203, A\*0204, A\*0205, A\*0206, A\*0207, A\*6802, and A\*6901. Finally, the B7-like supertype-confirmed alleles are: B7, B\*3501-03, B51, B\*5301, B\*5401, B\*5501-2, B\*5601, B\*6701, and B\*7801 (potentially also B\*1401, B\*3504-06, B\*4201, and B\*5602).

- 5 Population coverage achieved by combining the A2-, A3- and B7-supertypes is approximately 86% in five major ethnic groups (see Table XXI). Coverage may be extended by including peptides bearing the A1 and A24 motifs. On average, A1 is present in 12% and A24 in 29% of the population across five different major ethnic groups (Caucasian, North American Black, Chinese, Japanese, and Hispanic). Together, 10 these alleles are represented with an average frequency of 39% in these same ethnic populations. The total coverage across the major ethnicities when A1 and A24 are combined with the coverage of the A2-, A3- and B7-supertype alleles is >95%.

Population coverage for HLA class II molecules can be developed analogously based on the present disclosure.

15

#### *Summary of candidate HLA class I and class II epitopes*

- In summary, on the basis of the data presented above, 34 conserved CTL epitopes were selected as vaccine candidates (Table XXXVII). Of these 34 epitopes, 7 are derived from core, 18 from polymerase, and 9 from envelope. No epitopes from the X antigen 20 were included in the package as this protein is expressed in low amounts and is, therefore, of less immunological interest.

- The population coverage afforded by this panel of CTL epitopes is estimated to exceed 95% in each of 5 major ethnic populations. Using a Monte Carlo analysis (Figure 1), it is predicted that approximately 90% of the individuals in a population comprised of 25 Caucasians, North American Blacks, Japanese, Chinese and Hispanics would recognize five or more of the vaccine candidate epitopes.

- While preferred CTL epitopes includes 34 discrete peptides, two peptides are entirely nested within longer peptides, thus effectively reducing the numbers of peptides that would have to be included in a vaccine candidate. Specifically, the A2-restricted peptide 927.15 is nested within the B7-restricted peptide 26.0570 and the B7-restricted peptide 988.05 is nested within the A2-restricted peptide 924.07. Similarly, the A24-restricted peptide 20.0136 and the A2-restricted peptide 1013.01 contain the same core region, differing only at the first amino acid. On a related note, the A2-restricted peptide 30



1090.14 and the B7-restricted peptide 1147.05 overlap by two amino acids, raising the possibility of delivering these two epitopes as one contiguous peptide sequence.

The set of recommended vaccine candidates includes 9 A2-restricted CTL epitopes; four polymerase-derived epitopes, four envelope-derived epitopes and a core epitope. Seven of these 9 peptides are recognized in recall CTL assays from acute patients. Of the 7 peptides recognized in patients, 2 are non-crossreactive binding peptides. The inclusion of these peptides as potential vaccine candidates stems from the observation that HLA-A\*0201 is the predominantly expressed A2-supertype allele in all ethnicities examined. As such, inclusion of non-crossreactive A\*0201 binding peptides increases the redundancy of antigen coverage and population coverage. The only two A2-restricted peptides that lack patient immunogenicity data are peptides 1090.77 and 1069.06. The 1090.77 peptide is an analog of a highly immunogenic peptide recognized in acute HBV patients. Although recall responses in patients have not been tested for the ability to recognize the analog peptide, immunogenicity studies conducted in HLA transgenic mice have shown that CTL induced with 1090.77 are capable of recognizing target cells loaded with the naturally occurring sequence. This data indicates that CTL raised to the 1090.77 peptide are cross-reactive and should recognize HBV-infected cells. The 1069.06 peptide was included as a potential vaccine epitope because its high binding affinity for A\*6802 results in a greater population coverage. While peptide 1069.06 has not been tested for recognition by acute HBV patients, the peptide is immunogenic in HLA-A2 transgenic mice and primary human cultures.

Preferred CTL epitopes include 7 A3-supertype-restricted peptides; 6 derived from the polymerase antigen, and one from the core region. All of the A3-supertype vaccine candidate peptides are immunogenic in patients. Although peptide 1142.05 is a non-crossreactive A3-restricted peptide, it has been included because it has been shown to be recognized in patients and is capable of binding HLA-A1.

Nine B7-restricted peptides are preferred CTL epitopes. Of this group, 3 epitopes have been shown to be recognized in patients. While one of these peptides, 1147.04, is a non-crossreactive binder, it binds 2 of the major B7 supertype alleles with an  $IC_{50}$  or binding affinity value of less than 100 nM. Six B7-supertype epitopes were included as preferred epitopes based on supertype binding. Immunogenicity studies in humans (Bertoni et al., 1997; Doolan et al., 1997; Threlkeld et al., 1997) have demonstrated that highly cross-reactive peptides are almost always recognized as epitopes. Given these

results, and in light of the limited immunogenicity data available, the use of B7-supertype binding affinity as a selection criterion was deemed appropriate.

Similarly, there is little immunogenicity data regarding A1- and A24-restricted peptides. One preferred CTL epitope, 1069.04, has been reported to be recognized in recall responses from acute HBV patients. As discussed in the preceding paragraph, a high percentage of the peptides with binding affinities <100 nM are found to be immunogenic. For this reason, all A1 and A24 peptides with binding affinities <100 nM were considered as preferred CTL epitopes. Using this selection criterion, 3 A1-restricted and 6 A24-restricted peptides are identified as candidate epitopes. Further analysis found that 3 core-derived peptides bound A24 with intermediate affinity. Since relatively few core epitopes were identified during the course of this study, the intermediate A24 binding core peptides were also included in the set of preferred epitopes to provide a greater degree of redundancy in antigen coverage.

The list of preferred HBV-derived HTL epitopes is summarized in Table XXXVII. The set of HTL epitopes includes 12 DR supermotif binding peptides and 4 DR3 binding peptides. The bulk of the HTL epitopes are derived from polymerase; 2 envelope and 2 core derived epitopes are also included in the set of preferred HTL epitopes. The total estimated population coverage represented by the panel of HTL epitopes is in excess of 91% in each of five major ethnic groups (Table XXXVIII)

#### Example 7: Recognition Of Generation Of Endogenous Processed Antigens After Priming

This example determines that CTL induced by native or analogued peptide epitopes identified and selected as described in Examples 1-5 recognize endogenously synthesized, i.e., native antigens.

Effector cells isolated from transgenic mice that are immunized with peptide epitopes as in Example 3 are re-stimulated *in vitro* using peptide-coated stimulator cells. Six days later, effector cells are assayed for cytotoxicity and the cell lines that contain peptide-specific cytotoxic activity are further re-stimulated. An additional six days later, these cell lines are tested for cytotoxic activity on <sup>51</sup>Cr labeled 3A4-721.221-A11/K<sup>b</sup> target cells, in the absence or presence of peptide, and also tested on <sup>51</sup>Cr labeled target cells bearing the endogenously synthesized antigen, e.g., cells that are stably transfected with HBV expression vectors.

The result will demonstrate that CTL lines obtained from animals primed with peptide epitope recognize endogenously synthesized HBV antigen.

#### Example 8: Activity Of CTL-HTL Conjugated Epitopes In Transgenic Mice

This example illustrates the induction of CTLs in transgenic mice by use of an HBV CTL/HTL peptide conjugate. An analogous study may be found in Oseroff *et al. Vaccine* 16:823-833 (1998). The peptide composition can comprise multiple CTL and/or HTL epitopes. Such a peptide composition can comprise a lipidated HTL epitope conjugated to a preferred CTL epitope containing, for example, an A11 motif or an analog of that epitope.

Lipopeptides are prepared by coupling the appropriate fatty acid to the amino terminus of the resin bound peptide. A typical procedure is as follows: A dichloromethane solution of a four-fold excess of a pre-formed symmetrical anhydride of the appropriate fatty acid is added to the resin and the mixture is allowed to react for two hours. The resin is washed with dichloromethane and dried. The resin is then treated with trifluoroacetic acid in the presence of appropriate scavengers [e.g. 5% (v/v) water] for 60 minutes at 20°C. After evaporation of excess trifluoroacetic acid, the crude peptide is washed with diethyl ether, dissolved in methanol and precipitated by the addition of water. The peptide is collected by filtration and dried.

Preparation of peptides for immunization: Peptide compositions are typically resuspended in DMSO at a concentration of 20 mg/ml. Before use, peptides are prepared at the required concentration by dilution in saline or the appropriate medium.

Immunization procedures: A11/K<sup>b</sup> mice, which are transgenic for the human HLA A11 allele, are primed subcutaneously (base of the tail) with 0.1 ml of peptide conjugate formulated in saline, or DMSO/saline. Seven days after priming, splenocytes obtained from these animals are restimulated with syngeneic irradiated LPS-activated lymphoblasts coated with peptide.

#### Media:

a. RPMI-1640 supplemented with 10% fetal calf serum (FCS) 2 mM Glutamine, 50 µg/ml Gentamicin and  $5 \times 10^{-5}$  M 2-mercaptoethanol serves as culture medium

b. RPMI-1640 containing 25 mM HEPES buffer and supplemented with 2% (FCS) is used as cell washing medium.

Cell lines: The 3A4-721.221-A11/K<sup>b</sup> cell line is used as target cells. This cell line is an EBV transformed cell line that was mutagenized and selected to be Class I negative which was transfected with an HLA-A11/K<sup>b</sup> gene.

5 LPS-activated lymphoblasts: Splenocytes obtained from transgenic mice are resuspended at a concentration of  $1-1.5 \times 10^6$ /ml in culture medium supplemented with 25  $\mu$ g/ml LPS and 7  $\mu$ g/ml dextran sulfate in 75 cm<sup>2</sup> tissue culture flasks. After 72 hours at 37°C, the lymphoblasts are collected for use by centrifugation.

10 Peptide coating of lymphoblasts: Peptide coating of the LPS activated lymphoblasts is achieved by incubating  $30 \times 10^6$  irradiated (3000 rads) lymphoblasts with 100  $\mu$ g of peptide in 1 ml of R10 medium for 1 hr at 37°C. Cells are then washed once and resuspended in culture medium at the desired concentration.

15 *In vitro* CTL activation: One week after priming, spleen cells ( $30 \times 10^6$  cells/flask) are co-cultured at 37°C with syngeneic, irradiated (3000 rads), peptide coated lymphoblasts ( $10 \times 10^6$  cells/flask) in 10 ml of culture medium/T25 flask. After six days, the effector cells are harvested and assayed for cytotoxic activity.

20 Assay for cytotoxic activity: Target cells ( $1.0-1.5 \times 10^6$ ) are incubated at 37°C in the presence of 200  $\mu$ l of sodium <sup>51</sup>Cr chromate. After 60 minutes, cells are washed three times and resuspended in R10 medium. Peptide is added where required at a concentration of 1  $\mu$ g/ml. For the assay,  $104 \times 10^3$  <sup>51</sup>Cr-labeled target cells are added to different concentrations of effector cells (final volume of 200  $\mu$ l) in U-bottom 96-well plates. After a 6 hour incubation period at 37°C, a 0.1 ml aliquot of supernatant is removed from each well and radioactivity is determined in a Micromedic automatic gamma counter. The percent specific lysis is determined by the formula: percent specific release =  $100 \times (\text{experimental release} - \text{spontaneous release}) / (\text{maximum release} - \text{spontaneous release})$ . To facilitate comparison between separate CTL assays run under the same conditions, % <sup>51</sup>Cr release data is expressed as lytic units/ $10^6$  cells. One lytic unit is arbitrarily defined as the number of effector cells required to achieve 30% lysis of 10,000 target cells in a 6 hour <sup>51</sup>Cr release assay. To obtain specific lytic units/ $10^6$ , the lytic units/ $10^6$  obtained in the absence of peptide is subtracted from the lytic units/ $10^6$  obtained in the presence of peptide. For example, if 30% <sup>51</sup>Cr release is obtained at the E:T of 50:1 (i.e.,  $5 \times 10^5$  effector cells for 10,000 targets) in the absence of peptide and 5:1 (i.e.,  $5 \times 10^4$  effector cells for 10,000 targets) in the presence of peptide, the specific lytic units would be:  $(1 \times 10^6(5 \times 10^4)) - (1 \times 10^6(5 \times 10^5)) = 18 \text{LU}/10^6$ .

25

30

The results are analyzed to assess the magnitude of the CTL responses of animals injected with the immunogenic CTL/HTL conjugate vaccine preparation and are compared to the magnitude of the CTL response achieved using the CTL epitope as outlined in Example 3. Analyses similar to this may be performed to evaluate the immunogenicity of peptide conjugates containing multiple CTL epitopes and/or multiple HTL epitopes. In accordance with these procedures it is found that a CTL response is induced, and concomitantly that an HTL response is induced upon administration of such compositions.

10 Example 9. Selection of CTL and HTL epitopes for inclusion in an HBV-specific vaccine.

This example illustrates the procedure for the selection of peptide epitopes for vaccine compositions of the invention.

15 The following principles are utilized when selecting an array of epitopes for inclusion in a polypeptidic composition, or for selecting epitopes to be included in a vaccine composition and/or to be encoded by a minigene. Each of the following principles are balanced in order to make the selection.

20 1.) Epitopes are selected which, upon administration, mimic immune responses that have been observed to be correlated with HBV clearance. For HLA Class I this includes 3-4 epitopes that come from at least one antigen of HBV. In other words, it has been observed that in patients who spontaneously clear HBV, that they had generated an immune response to at least 3 epitopes on at least one HBV antigen. For HLA Class II a similar rationale is employed; again 3-4 epitopes are selected from at least one HBV antigen.

2.) Epitopes are selected that have the requisite binding affinity established to be correlated with immunogenicity: for HLA Class I an  $IC_{50}$  of 500 nM or less, or for Class II an  $IC_{50}$  of 1000 nM or less.

30 3.) Sufficient supermotif bearing peptides, or a sufficient array of allele-specific motif bearing peptides, are selected to give broad population coverage. For example, epitopes are selected to provide at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, is employed to assess population coverage.

4.) When selecting epitopes for HBV antigens it is often preferable to select native epitopes. Therefore, of particular relevance for infectious disease vaccines, are epitopes referred to as “nested epitopes.” Nested epitopes occur where at least two epitopes overlap in a given peptide sequence. A peptide comprising “transcendent nested  
5 epitopes” is a peptide that has both HLA class I and HLA class II epitopes in it.

When providing nested epitopes, a sequence that has the greatest number of epitopes per provided sequence is provided. A limitation on this principle is to avoid providing a peptide that is any longer than the amino terminus of the amino terminal epitope and the carboxyl terminus of the carboxyl terminal epitope in the peptide. When  
10 providing a longer peptide sequence, such as a sequence comprising nested epitopes, the sequence is screened in order to insure that it does not have pathological or other deleterious biological properties.

5.) When creating a minigene, as disclosed in greater detail in the Example 9, an objective is to generate the smallest peptide possible that encompasses the epitopes of  
15 interest. The principles employed are similar, if not the same as those employed when selecting a peptide comprising nested epitopes. Thus, upon determination of the nucleic acid sequence to be provided as a minigene, the peptide encoded thereby is analyzed to determine whether any “junctional epitopes” have been created. A junctional epitope is an actual binding epitope, as predicted, e.g., by motif analysis. Junctional epitopes are to  
20 be avoided because the recipient may generate an immune response to that epitope. Of particular concern is a junctional epitope that is a “dominant epitope.” A dominant epitope may lead to such a zealous response that immune responses to other epitopes are diminished or suppressed.

Peptide epitopes for inclusion in vaccine compositions are, for example, selected  
25 from those listed in Table XXXVIIa and b. A vaccine composition comprised of selected peptides, when administered, is safe, efficacious, and elicits an immune response similar in magnitude of an immune response that clears an acute HBV infection.

#### Example 10: Construction of Minigene Multi-Epitope DNA Plasmids

30 This example provides an illustration of the construction of a minigene expression plasmid. Minigene plasmids may, of course, contain various configurations of CTL and/or HTL epitopes or epitope analogs as described herein. Expression plasmids have been constructed and evaluated as described, for example, in U.S.S.N. 60/085,751 filed 5/15/98 and U.S.S.N. 09/078,904 filed 5/13/98. An example of such a plasmid is shown

in Figure 2, which illustrates the orientation of HBV epitopes in minigene constructs. Such a plasmid may, for example, also include multiple CTL and HTL peptide epitopes.

A minigene expression plasmid may include multiple CTL and HTL peptide epitopes. In the present example, HLA-A2, -A3, -B7 supermotif-bearing peptide epitopes and HLA-A1 and -A24 motif-bearing peptide epitopes are used in conjunction with DR supermotif-bearing epitopes and/or DR3 epitopes (Figure 2). Preferred epitopes are identified, for example, in Tables XXVI-XXXIII, HLA class I supermotif or motif-bearing peptide epitopes derived from multiple HBV antigens, *e.g.*, the core, polymerase, envelope and X proteins, are selected such that multiple supermotifs/motifs are represented to ensure broad population coverage. Similarly, HLA class II epitopes are selected from multiple HBV antigens to provide broad population coverage, *i.e.* both HLA DR-1-4-7 supermotif-bearing epitopes and HLA DR-3 motif-bearing epitopes are selected for inclusion in the minigene construct. The selected CTL and HTL epitopes are then incorporated into a minigene for expression in an expression vector.

This example illustrates the methods to be used for construction of such a minigene-bearing expression plasmid. Other expression vectors that may be used for minigene compositions are available and known to those of skill in the art.

The minigene DNA plasmid contains a consensus Kozak sequence and a consensus murine kappa Ig-light chain signal sequence followed by a string of CTL and/or HTL epitopes selected in accordance with principles disclosed herein.

Overlapping oligonucleotides, for example eight oligonucleotides, averaging approximately 70 nucleotides in length with 15 nucleotide overlaps, are synthesized and HPLC-purified. The oligonucleotides encode the selected peptide epitopes as well as appropriate linker nucleotides. The final multiepitope minigene is assembled by extending the overlapping oligonucleotides in three sets of reactions using PCR. A Perkin/Elmer 9600 PCR machine is used and a total of 30 cycles are performed using the following conditions: 95°C for 15 sec, annealing temperature (5° below the lowest calculated T<sub>m</sub> of each primer pair) for 30 sec, and 72°C for 1 min.

For the first PCR reaction, 5 µg of each of two oligonucleotides are annealed and extended: Oligonucleotides 1+2, 3+4, 5+6, and 7+8 are combined in 100 µl reactions containing *Pfu* polymerase buffer (1x= 10 mM KCL, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 20 mM Tris-chloride, pH 8.75, 2 mM MgSO<sub>4</sub>, 0.1% Triton X-100, 100 µg/ml BSA), 0.25 mM each dNTP, and 2.5 U of *Pfu* polymerase. The full-length dimer products are gel-purified, and two reactions containing the product of 1+2 and 3+4, and the product of 5+6 and 7+8 are

mixed, annealed, and extended for 10 cycles. Half of the two reactions are then mixed, and 5 cycles of annealing and extension carried out before flanking primers are added to amplify the full length product for 25 additional cycles. The full-length product is gel-purified and cloned into pCR-blunt (Invitrogen) and individual clones are screened by sequencing.

Example 11. The plasmid construct and the degree to which it induces immunogenicity.

The degree to which the plasmid construct prepared using the methodology outlined in Example 9 is able to induce immunogenicity is evaluated through *in vivo* injections into transgenic mice and *in vitro* culture of CTL and HTL, which are subsequently analysed using cytotoxicity and proliferation assays, respectively, as detailed *e.g.*, in U.S.S.N. 09/311,784 filed 5/13/99 and Alexander *et al.*, *Immunity* 1:751-761, 1994. To assess the capacity of the pMin minigene construct to induce CTLs *in vivo*, HLA-A11/K<sup>b</sup> transgenic mice, for example, are immunized intramuscularly with 100 µg of plasmid cDNA. As a means of comparing the level of CTLs induced by DNA immunization, a control group of animals is also immunized with an actual peptide composition that comprises multiple epitopes synthesized as a single polypeptide as they would be encoded by the minigene.

Splenocytes from immunized animals are stimulated twice with each of the respective compositions (peptide epitopes encoded in the minigene or the polyepitopic peptide), then assayed for peptide-specific cytotoxic activity in a <sup>51</sup>Cr release assay. The results indicate the magnitude of the CTL response directed against the A3-restricted epitope, thus indicating the *in vivo* immunogenicity of the minigene vaccine and polyepitopic vaccine. It is, therefore, found that the minigene elicits immune responses directed toward the HLA-A3 supermotif peptide epitopes as does the polyepitopic peptide vaccine. Such an analysis is also performed using other HLA-A2 and HLA-B7 transgenic mouse models to assess CTL induction by HLA-A2 and HLA-B7 motif or supermotif epitopes.

To assess the capacity of a class II epitope encoding minigene to induce HTLs *in vivo*, I-A<sup>b</sup> restricted mice, for example, are immunized intramuscularly with 100 µg of plasmid DNA. As a means of comparing the level of HTLs induced by DNA immunization, a group of control animals is also immunized with an actual peptide composition emulsified in complete Freund's adjuvant.



CD4+ T cells, *i.e.* HTLs, are purified from splenocytes of immunized animals and stimulated with each of the respective compositions (peptides encoded in the minigene). The HTL response is measured using a  $^3\text{H}$ -thymidine incorporation proliferation assay, (see, *e.g.*, Alexander et al. Immunity 1:751-761, 1994). the results indicate the magnitude  
 5 of the HTL response , thus demonstrating the *in vivo* immunogenicity of the minigene.

#### Example 12: Peptide Composition for Prophylactic Uses

Vaccine compositions of the present invention are used to prevent HBV infection in persons who are at risk for such an infection. For example, a polyepitopic peptide  
 10 epitope composition containing multiple CTL and HTL epitopes such as those selected in Examples 9 and/or 10, which are also selected to target greater than 80% of the population, is administered to individuals at risk for HBV infection. The composition is provided as a single lipidated polypeptide that encompasses multiple epitopes. The vaccine is administered in an aqueous carrier comprised of Freund's Incomplete Adjuvant.  
 15 The dose of peptide for the initial immunization is from about 500 to about 50,000  $\mu\text{g}$  for a 70 kg patient. The initial administration of vaccine is followed by booster dosages at 4 weeks followed by evaluation of the magnitude of the immune response in the patient by techniques that determine the presence of epitope-specific CTL populations in a PBMC sample. Additional booster doses are administered as required. The composition is found  
 20 to be both safe and efficacious as a prophylaxis against HBV infection.

Alternatively, the polyepitopic peptide composition can be administered as a nucleic acid in accordance with methodologies known in the art and disclosed herein.

#### Example 13: Polyepitopic Vaccine Compositions Derived from Native HBV Sequences

25 A native HBV polyprotein sequence is screened, preferably using computer algorithms defined for each class I and/or class II supermotif or motif, to identify "relatively short" regions of the polyprotein that comprise multiple epitopes. This relatively short sequence that contains multiple distinct, even overlapping, epitopes is selected and used to generate a minigene construct. The construct is engineered to  
 30 express the peptide, which corresponds to the native protein sequence. The "relatively short" peptide is less than 250 amino acids in length, preferably less than 100 amino acids in length, and more preferably less than 75 or 50 amino acids in length. The protein sequence of the vaccine composition is selected because it has maximal number of epitopes contained within the sequence. As noted herein, epitope motifs may be

overlapping (*i.e.*, frame shifted relative to one another) with frame shifted overlapping epitopes, *e.g.* two 9-mer epitopes can be present in a 10 amino acid peptide. Such a vaccine composition is administered for therapeutic or prophylactic purposes.

The vaccine composition will preferably include, for example, three CTL epitopes and at least one HTL epitope from the source antigen. This polyepitopic native sequence is administered either as a peptide or as a nucleic acid sequence which encodes the peptide. Alternatively, an analog can be made of this native sequence, whereby one or more of the epitopes comprise substitutions that alter the cross-reactivity and/or binding affinity properties of the polyepitopic peptide.

The embodiment of this example provides for the possibility that an as yet undiscovered aspect of immune system processing will apply to the native nested sequence and thereby facilitate the production of therapeutic or prophylactic immune response-inducing vaccine compositions. Additionally such an embodiment provides for the possibility of motif-bearing epitopes for an HLA makeup that is presently unknown. Furthermore, this embodiment (absent analogs) directs the immune response to peptide sequences that are present in native HBV antigens. Lastly, the embodiment provides an economy of scale when producing nucleic acid vaccine compositions.

Related to this embodiment, computer programs can be derived which identify, in a target sequence, the greatest number of epitopes per sequence length.

#### Example 14. Polyepitopic Vaccine Compositions Directed To Multiple Diseases

The HBV peptide epitopes of the present invention are used in conjunction with peptide epitopes from target antigens related to one or more other diseases, to create a vaccine composition that is useful for the prevention or treatment of HBV as well as another disease. Examples of other diseases include, but are not limited to, HIV, HCV, and HPV.

For example, a polyepitopic peptide composition comprising multiple CTL and HTL epitopes that target greater than 98% of the population may be created for administration to individuals at risk for both HBV and HIV infection. The composition can be provided as a single polypeptide that incorporates the multiple epitopes from the various disease-associated sources, or can be administered as a composition comprising one or more discrete epitopes.

### Example 15. Use of peptides to evaluate an immune response

Peptides of the invention may be used to analyze an immune response for the presence of specific CTL populations corresponding to HBV. Such an analysis may be performed as described by Ogg *et al.*, *Science* 279:2103-2106, 1998. In the following example, peptides in accordance with the invention are used as a reagent for diagnostic or prognostic purposes, not as an immunogen.

In this example highly sensitive human leukocyte antigen tetrameric complexes ("tetramers") may be used for a cross-sectional analysis of, for example, HBV Env-specific CTL frequencies from untreated HLA A\*0201-positive individuals at different stages of infection using an HBV Env peptide containing an A2.1 extended motif. Tetrameric complexes are synthesized as described (Musey *et al.*, *N. Engl. J. Med.* 337:1267, 1997). Briefly, purified HLA heavy chain (A2.1 in this example) and  $\beta$ 2-microglobulin are synthesized by means of a prokaryotic expression system. The heavy chain is modified by deletion of the transmembrane-cytosolic tail and COOH-terminal addition of a sequence containing a BirA enzymatic biotinylation site. The heavy chain,  $\beta$ 2-microglobulin, and peptide are refolded by dilution. The 45-kD refolded product is isolated by fast protein liquid chromatography and then biotinylated by BirA in the presence of biotin (Sigma, St. Louis, Missouri), adenosine 5'triphosphate and magnesium. Streptavidin-phycoerythrin conjugate is added in a 1:4 molar ratio, and the tetrameric product is concentrated to 1 mg/ml. The resulting product is referred to as tetramer-phycoerythrin.

For the analysis of patient blood samples, approximately one million PBMCs are centrifuged at 300g for 5 minutes and resuspended in 50 ul of cold phosphate-buffered saline. Tri-color analysis is performed with the tetramer-phycoerythrin, along with anti-CD8-Tricolor, and anti-CD38. The PBMCs are incubated with tetramer and antibodies on ice for 30 to 60 min and then washed twice before formaldehyde fixation. Gates are applied to contain >99.98% of control samples. Controls for the tetramers include both A\*0201-negative individuals and A\*0201-positive uninfected donors. The percentage of cells stained with the tetramer is then determined by flow cytometry. The results indicate the number of cells in the PBMC sample that contain epitope-restricted CTLs, thereby readily indicating the stage of infection with HBV or the status of exposure to HBV or to a vaccine that elicits a protective response.

### Example 16: Use of Peptide Epitopes to Evaluate Recall Responses

The peptide epitopes of the invention are used as reagents to evaluate T cell responses such as acute or recall responses, in patients. Such an analysis may be performed on patients who have recovered from infection or who are chronically infected with HBV or who have been vaccinated with an HBV vaccine.

For example, the class I restricted CTL response of persons at risk for HBV infection who have been vaccinated may be analyzed. The vaccine may be any HBV vaccine. PBMC are collected from vaccinated individuals and HLA typed. Appropriate peptide reagents that, are highly conserved and, optimally, bear supermotifs to provide cross-reactivity with multiple HLA supertype family members are then used for analysis of samples derived from individuals who bear that HLA type.

PBMC from vaccinated individuals are separated on Ficoll-Histopaque density gradients (Sigma Chemical Co., St. Louis, MO), washed three times in HBSS (GIBCO Laboratories), resuspended in RPMI-1640 (GIBCO Laboratories) supplemented with L-glutamine (2mM), penicillin (50U/ml), streptomycin (50 µg/ml), and Hepes (10mM) containing 10% heat-inactivated human AB serum (complete RPMI) and plated using microculture formats. Synthetic peptide is added at 10 µg/ml to each well and recombinant HBc Ag is added at 1 µg/ml to each well as a source of T cell help during the first week of stimulation.

In the microculture format,  $4 \times 10^5$  PBMC are stimulated with peptide in 8 replicate cultures in 96-well round bottom plate in 100 µl/well of complete RPMI. On days 3 and 10, 100 ml of complete RPMI and 20 U/ml final concentration of rIL-2 are added to each well. On day 7 the cultures are transferred into a 96-well flat-bottom plate and restimulated with peptide, rIL-2 and  $10^5$  irradiated (3,000 rad) autologous feeder cells. The cultures are tested for cytotoxic activity on day 14. A positive CTL response requires two or more of the eight replicate cultures to display greater than 10% specific  $^{51}\text{Cr}$  release, based on comparison with uninfected control subjects as previously described (Rehermann, *et al.*, *Nature Med.* 2:1104,1108, 1996; Rehermann *et al.*, *J. Clin. Invest.* 97:1655-1665, 1996; and Rehermann *et al.* *J. Clin. Invest.* 98:1432-1440, 1996).

Target cell lines are autologous and allogeneic EBV-transformed B-LCL that are either purchased from the American Society for Histocompatibility and Immunogenetics (ASHI, Boston, MA) or established from the pool of patients as described (Guilhot, *et al.* *J. Virol.* 66:2670-2678, 1992).

Cytotoxicity assays are performed in the following manner. Target cells consist of either allogeneic HLA-matched or autologous EBV-transformed B lymphoblastoid cell line that are incubated overnight with synthetic peptide at 10  $\mu$ M and labeled with 100  $\mu$ Ci of  $^{51}$ Cr (Amersham Corp., Arlington Heights, IL) for 1 hour after which they are washed four times with HBSS. Cytolytic activity is determined in a standard 4-h, split well  $^{51}$ Cr release assay using U-bottomed 96 well plates containing 3,000 targets/well. Stimulated PBMC are tested at E/T ratios of 20-50:1 on day 14. Percent cytotoxicity is determined from the formula:  $100 \times [(\text{experimental release} - \text{spontaneous release}) / (\text{maximum release} - \text{spontaneous release})]$ . Maximum release is determined by lysis of targets by detergent (2% Triton X-100; Sigma Chemical Co., St. Louis, MO). Spontaneous release is <25% of maximum release for all experiments.

The results of such an analysis will indicate to what extent HLA-restricted CTL populations have been stimulated with the vaccine. Of course, this protocol can also be used to monitor prior HBV exposure.

The class II restricted HTL responses may also be analyzed. Purified PBMC are cultured in a 96-well flat bottom plate at a density of  $1.5 \times 10^5$  cells/well and are stimulated with 10  $\mu$ g/ml synthetic peptide, whole antigen, or PHA. Cells are routinely plated in replicates of 4-6 wells for each condition. After seven days of culture, the medium is removed and replaced with fresh medium containing 10U/ml IL-2. Two days later, 1  $\mu$ Ci  $^3$ H-thymidine is added to each well and incubation is continued for an additional 18 hours. Cellular DNA is then harvested on glass fiber mats and analyzed for  $^3$ H-thymidine incorporation. Antigen-specific T cell proliferation is calculated as the ratio of  $^3$ H-thymidine incorporation in the presence of antigen divided by the  $^3$ H-thymidine incorporation in the absence of antigen.

The results of such an analysis will indicate to what extent HLA-restricted HTL populations have been stimulated with a vaccine or prior exposure to HBV.

#### Example 17: Induction Of Specific CTL Response In Humans

A human clinical trial for an immunogenic composition comprising HBV CTL and HTL epitopes of the invention is set up as an IND Phase I, dose escalation study (5, 50 and 500  $\mu$ g) and carried out as a randomized, double-blind, placebo-controlled trial. Such a trial is designed, for example, as follows:

A total of about 27 subjects are enrolled and divided into 3 groups:

Group I: 3 subjects are injected with placebo and 6 subjects are injected with 5 µg of peptide composition;

Group II: 3 subjects are injected with placebo and 6 subjects are injected with 50 µg peptide composition;

5       Group III: 3 subjects are injected with placebo and 6 subjects are injected with 500 µg of peptide composition.

After 4 weeks following the first injection, all subjects receive a booster inoculation at the same dosage.

10       The endpoints measured in this study relate to the safety and tolerability of the peptide composition as well as its immunogenicity. Cellular immune responses to the peptide composition are an index of the intrinsic activity of this the peptide composition, and can therefore be viewed as a measure of biological efficacy. The following summarize the clinical and laboratory data that relate to safety and efficacy endpoints.

15       Safety: The incidence of adverse events is monitored in the placebo and drug treatment group and assessed in terms of degree and reversibility.

Evaluation of Vaccine Efficacy: For evaluation of vaccine efficacy, subjects are bled before and after injection. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

20       Thus, the vaccine is found to be both safe and efficacious.

#### Example 18: Phase II Trials In Patients Infected With HBV

25       Phase II trials are performed to study the effect of administering the CTL-HTL peptide compositions to patients (male and female ) having chronic HBV infection. A main objective of the trials is to determine an effective dose and regimen for inducing CTLs in chronically infected HBV patients, to establish the safety of inducing a CTL response in these patients, and to see to what extent activation of CTLs improves the clinical picture of chronically infected CTL patients, as manifested by a transient flare in alanine aminotransferase (ALT), normalization of ALT, and reduction in HBV DNA.

30       Such a study is designed, for example, as follows:

The studies are performed in multiple centers in the U.S. and Canada. The trial design is an open-label, uncontrolled, dose escalation protocol wherein the peptide composition is administered as a single dose followed six weeks later by a single booster

shot of the same dose. The dosages are 50, 500 and 5,000 micrograms per injection. Drug-associated adverse effects are recorded.

There are three patient groupings. The first group is injected with 50 micrograms of the peptide composition and the second and third groups with 500 and 5,000  
5 micrograms of peptide composition, respectively. The patients within each group range in age from 21-65 and include both males and females. The patients represent diverse ethnic backgrounds. All of them are infected with HBV for over five years and are HIV, HCV and HDV negative, but have positive levels of HBe antigen and HBs antigen.

The magnitude and incidence of ALT flares and the levels of HBV DNA in the  
10 blood are monitored to assess the effects of administering the peptide compositions. The levels of HBV DNA in the blood are an indirect indication of the progress of treatment. The vaccine composition is found to be both safe and efficacious in the treatment of chronic HBV infection.

The examples herein are provided to illustrate the invention but not to limit its  
15 scope. For example, the human terminology for the Major Histocompatibility Complex, namely HLA, is used throughout this document. It is to be appreciated that these principles can be extended to other species as well. Moreover, peptide epitopes have been disclosed in the related application U.S.S.N. 08/820,360, which was previously incorporated by reference. Thus, other variants of the invention will be readily apparent  
20 to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, and patent application cited herein are hereby incorporated by reference for all purposes.

TABLE I

SUPERMOTIFS	POSITION	POSITION	POSITION
	2 (Primary Anchor)	3 (Primary Anchor)	C Terminus (Primary Anchor)
A1	<b>T</b> <i>L</i> <b>V</b> <i>M</i> <b>S</b>		<b>F</b> <i>W</i> <b>Y</b>
A2	<b>L</b> <i>I</i> <b>V</b> <i>M</i> <b>A</b> <i>T</i> <b>Q</b>		<b>I</b> <i>V</i> <b>M</b> <i>A</i> <b>T</b> <i>L</i>
A3	<b>V</b> <i>S</i> <b>M</b> <i>A</i> <b>T</b> <i>L</i>		<b>R</b> <b>K</b>
A24	<b>Y</b> <i>F</i> <b>W</b> <i>I</i> <b>V</b> <i>L</i> <b>M</b> <i>T</i>		<b>F</b> <i>I</i> <b>Y</b> <i>W</i> <b>L</b> <i>M</i>
B7	<b>P</b>		<b>V</b> <i>I</i> <b>L</b> <i>F</i> <b>M</b> <i>W</i> <b>Y</b> <i>A</i>
B27	<b>R</b> <b>H</b> <b>K</b>		<b>F</b> <i>Y</i> <b>L</b> <i>W</i> <b>M</b> <i>I<b>V</b><i>A</i></i>
B44	<b>E</b> <b>D</b>		<b>F</b> <i>W</i> <b>Y</b> <i>L</i> <b>I</b> <b>M</b> <i>V<b>A</b></i>
B58	<b>A</b> <b>T</b> <b>S</b>		<b>F</b> <i>W</i> <b>Y</b> <i>L</i> <b>I</b> <b>V</b> <i>M<b>A</b></i>
B62	<b>Q</b> <i>L</i> <b>I</b> <b>V</b> <i>M<b>P</b></i>		<b>F</b> <i>W</i> <b>Y</b> <i>M</i> <b>I</b> <b>V</b> <i>L<b>A</b></i>
MOTIFS			
A1	<b>T</b> <b>S</b> <b>M</b>		<b>Y</b>
A1		<b>D</b> <b>E</b> <b>A</b> <b>S</b>	<b>Y</b>
A2.1	<b>L</b> <i>M</i> <b>V</b> <i>Q</i> <b>I</b> <b>A</b> <i>T</i>		<b>V</b> <i>L</i> <b>I</b> <b>M</b> <i>A<b>T</b></i>
A3	<b>L</b> <i>M</i> <b>V</b> <i>I<b>S</b><i>A</i><b>T</b><i>F</i><b>C</b><i>G<b>D</b></i></i>		<b>K</b> <i>Y</i> <b>R</b> <i>H</i> <b>F</b> <i>A</i>
A11	<b>V</b> <i>T</i> <b>M</b> <i>L</i> <b>I</b> <b>S</b> <i>A</i> <b>G</b> <i>N<b>C</b><i>D<b>F</b></i></i>		<b>K</b> <i>R</i> <b>Y</b> <i>H</i>
A24	<b>Y</b> <b>F</b> <b>W</b> <b>M</b>		<b>F</b> <b>L</b> <b>I</b> <b>W</b>
A*3101	<b>M</b> <b>V</b> <i>T</i> <b>A</b> <b>L</b> <b>I</b> <b>S</b>		<b>R</b> <b>K</b>
A*3301	<b>M</b> <b>V</b> <i>A<b>L</b><i>F</i><b>I</b><b>S</b><b>T</b></i>		<b>R</b> <b>K</b>
A*6801	<b>A</b> <b>V</b> <i>T</i> <b>M</b> <b>S</b> <b>L</b> <b>I</b>		<b>R</b> <b>K</b>
B*0702	<b>P</b>		<b>L</b> <i>M</i> <b>F</b> <i>W</i> <b>Y</b> <i>A</i> <b>I</b> <b>V</b>
B*3501	<b>P</b>		<b>L</b> <i>M</i> <b>F</b> <i>W</i> <b>Y</b> <i>I<b>V</b><b>A</b></i>
B51	<b>P</b>		<b>L</b> <i>I</i> <b>V</b> <i>F</i> <b>W</b> <i>Y<b>A</b><b>M</b></i>
B*5301	<b>P</b>		<b>I</b> <i>M</i> <b>F</b> <i>W</i> <b>Y</b> <i>A<b>L</b><b>V</b></i>
B*5401	<b>P</b>		<b>A</b> <i>T</i> <b>I</b> <b>V</b> <i>L</i> <b>M</b> <i>F</i> <b>W</b> <i>Y</i>

Bolded residues are preferred, italicized residues are less preferred: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.



TABLE Ia

SUPERMOTIFS	POSITION	POSITION	POSITION
	2 (Primary Anchor)	3 (Primary Anchor)	C Terminus (Primary Anchor)
A1	<b>T</b> <i>ILVMS</i>		<b>F</b> <i>WY</i>
A2	<i>VQAT</i>		<i>VLIMAT</i>
A3	<b>V</b> <i>SMATLI</i>		<b>R</b> <i>K</i>
A24	<b>Y</b> <i>FWIVLMT</i>		<b>F</b> <i>IYWLM</i>
B7	<b>P</b>		<b>V</b> <i>ILFMWYA</i>
B27	<b>R</b> <i>HK</i>		<b>F</b> <i>YLWMIVA</i>
B58	<b>A</b> <i>TS</i>		<b>F</b> <i>WYLIIVMA</i>
B62	<b>Q</b> <i>LIVMP</i>		<b>F</b> <i>WYMIIVLA</i>
MOTIFS			
A1	<b>T</b> <i>S</i> <b>M</b>		<b>Y</b>
A1		<b>D</b> <i>EAS</i>	<b>Y</b>
A2.1	<i>VQAT</i> *		<i>VLIMAT</i>
A3.2	<b>L</b> <i>MVISATFCGD</i>		<b>K</b> <i>YRHF</i>
A11	<b>V</b> <i>TMLISAGNCDF</i>		<b>K</b> <i>RH</i>
A24	<b>Y</b> <i>FW</i>		<b>F</b> <i>LIW</i>

\*If 2 is V, or Q, the C-term is not L

Bolded residues are preferred, italicized residues are less preferred: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

TABLE II

POSITION										
<div><div>1</div><div>2</div><div>3</div><div>4</div><div>5</div><div>6</div><div>7</div><div>8</div><div>C-terminus</div></div>										
SUPERMOTIFS										
A1		<div><div>1° Anchor</div><div>TLIVMS</div></div>								<div><div>1° Anchor</div><div>FWY</div></div>
A2		<div><div>1° Anchor</div><div>LIVMATQ</div></div>								<div><div>1° Anchor</div><div>LIVMAT</div></div>
A3	preferred	<div><div>1° Anchor</div><div>VSMATLI</div></div>		YFW (4/5)		YFW (3/5)		YFW (4/5)	P (4/5)	<div><div>1° Anchor</div><div>RK</div></div>
	deleterious	DE (3/5); P (5/5)		DE (4/5)						
A24		<div><div>1° Anchor</div><div>YFWIVLM</div><div>T</div></div>								<div><div>1° Anchor</div><div>FIYVLM</div></div>
B7	preferred	<div><div>FWY (5/5)</div><div>LIVM (3/5)</div></div>		<div><div>1° Anchor</div><div>P</div></div>		FWY (4/5)		<div><div>FWY (3/5)</div><div>1° Anchor</div><div>VILFMWYA</div></div>		
	deleterious	DE (3/5); P(5/5); G(4/5); A(3/5); QN (3/5)		DE (3/5)		G (4/5)	QN (4/5)	DE (4/5)		
B27		<div><div>1° Anchor</div><div>RHK</div></div>								<div><div>1° Anchor</div><div>FYLIMIVA</div></div>
B44		<div><div>1° Anchor</div><div>ED</div></div>								<div><div>1° Anchor</div><div>FWYLIIMVA</div></div>
B58		<div><div>1° Anchor</div><div>ATS</div></div>								<div><div>1° Anchor</div><div>FWYLIIVMA</div></div>
B62		<div><div>1° Anchor</div><div>QLIVMP</div></div>								<div><div>1° Anchor</div><div>FWYMIIVLA</div></div>

POSITION									
1									
2									
3									
4									
5									
6									
7									
8									
C-terminus									
MOTIFS									
A1 9-mer	preferred	GRHK	ASTCLIV M	1°Anchor DEAS	GSTC	ASTC	LIVM	DE	1°Anchor Y
	deleterious	A	RHKDEPY FW	DE	PQN	RHK	PG	GP	

POSITION									
	1	2	3	4	5	6	7	8	9 or C-terminus
A1 10-mer	preferred deleterious	YFW GP	<u>1°Anchor</u> STM	DEAQN M	A DE	YFWQN RHK	PASTC RHKYFW	GDE RHK	P A
A1 10-mer	preferred deleterious	YFW RHK	STCLVM RHKDEPY FW	<u>1°Anchor</u> DEAS	A P	YFW G	PG PRHK	G QN	YFW Y
A2.1 9-mer	preferred deleterious	YFW DEP	<u>1°Anchor</u> LMTVQAT	YFW DERKH	STC	YFW	A	P	<u>1°Anchor</u> VLIMAT
A2.1 10-mer	preferred deleterious	AYFW DEP	<u>1°Anchor</u> LMTVQAT	LVIM DE	G RKHA	P	RKH	FYWL DERK H	<u>1°Anchor</u> VLIMAT

POSITION

	1	2	3	4	5	6	7	8	9 or C-terminus
A3	preferred RHK	$\frac{1^\circ\text{Anchor}}{\text{LMVISAT}}\text{FCGD}$	YFW	PRHKYFW	A	YFW		P	$\frac{1^\circ\text{Anchor}}{\text{KYRHFA}}$
deleterious	DEP		DE						
A11	preferred A	$\frac{1^\circ\text{Anchor}}{\text{VTLMISA}}\text{GNCDF}$	YFW	YFW	A	YFW	YFW	P	$\frac{1^\circ\text{Anchor}}{\text{KRYH}}$
deleterious	DEP						A	G	
A24 9-mer	preferred YFWRHK	$\frac{1^\circ\text{Anchor}}{\text{YFWM}}$		STC			YFW	YFW	$\frac{1^\circ\text{Anchor}}{\text{FLIW}}$
deleterious	DEG		DE	G	QNP	DERHK	G	AQN	
A24 10-mer	preferred YFWM	$\frac{1^\circ\text{Anchor}}{\text{YFWM}}$		P	YFWP		P		$\frac{1^\circ\text{Anchor}}{\text{FLIW}}$
deleterious			GDE	QN	RHK	DE	A	QN	DEA
A3101	preferred RHK	$\frac{1^\circ\text{Anchor}}{\text{MVTALLS}}$	YFW	P		YFW	YFW	AP	$\frac{1^\circ\text{Anchor}}{\text{RK}}$
deleterious	DEP		DE	ADE	DE	DE	DE		

POSITION									
	1	2	3	4	5	6	7	8	9 C-terminus or C-terminus
A3301 preferred		<u>1°Anchor</u> MVALF/S T	YFW				A YFW		<u>1°Anchor</u> RK
deleterious	GP		DE						
A6801 preferred	YFWSTC	<u>1°Anchor</u> AVTMSLI			YFWLIIV M		YFW	P	<u>1°Anchor</u> RK
deleterious	GP		DEG		RHK			A	
B0702 preferred	RHKFWY	<u>1°Anchor</u> P	RHK		RHK	RHK	RHK	PA	<u>1°Anchor</u> LMFWYIIV
deleterious	DEQNP		DEP	DE	DE	GDE	QN	DE	
B3501 preferred	FWYLIIVM	<u>1°Anchor</u> P	FWY				FWY		<u>1°Anchor</u> LMFWYIIV
deleterious	AGP			G	G				
B51 preferred	LIVMFY	<u>1°Anchor</u> P	FWY	STC	FWY		G	FWY	<u>1°Anchor</u> LIVFWYAM
deleterious	AGPDERHKSTC				DE	G	DEQN	GDE	

POSITION								
	1	2	3	4	5	6	7	8
								9 C-terminus or C-terminus
B5301 preferred	LIVMF <sup>1°Anchor</sup> WY	P	FWY	STC	FWY	LIVMF <sup>1°Anchor</sup> WY	FWY	IMFWY <sup>1°Anchor</sup> ALV
deleterious	AGPQN					G	RHKQN	DE
B5401 preferred	FWY							
		L <sup>1°Anchor</sup> P	FWYLVIM		LVIM		ALIVM	FWYAP
								L <sup>1°Anchor</sup> ATVLMFWY
deleterious	GPQNDE		GDESTC		RHKDE	DE	QNDGE	DE

Italicized residues indicate less preferred or “tolerated” residues.  
The information in Table II is specific for 9-mers unless otherwise specified.

TABLE III

		POSITION								
<u>MOTIFS</u>		<u>1° anchor 1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>1° anchor 6</u>	<u>7</u>	<u>8</u>	<u>9</u>
DR4	preferred	FM $\overline{Y}$ L $\overline{I}$ $\overline{W}$	M	T		I	VSTCPAL $\overline{I}$ $\overline{M}$	MH		MH
	deleterious				W			R		WDE
DR1	preferred	MFL $\overline{I}$ $\overline{W}$ $\overline{Y}$			PAMQ		VMATSP $\overline{L}$ $\overline{I}$ $\overline{C}$	M		AVM
	deleterious		C	CH	FD	CWD		GDE	D	
DR7	preferred	MFL $\overline{I}$ $\overline{W}$ $\overline{Y}$	M	W	A		IVMSACT $\overline{P}$ $\overline{L}$	M		IV
	deleterious		C		G			GRD	N	G
DR Supermotif		MFL $\overline{I}$ $\overline{W}$ $\overline{Y}$					VMSTAC $\overline{P}$ $\overline{L}$ $\overline{I}$			

<u>DR3 MOTIFS</u>		<u>1° anchor 1</u>	<u>2</u>	<u>3</u>	<u>1° anchor 4</u>	<u>5</u>	<u>1° anchor 6</u>
motif a							
preferred		LIVMFY			D		
motif b							
preferred		LIVMFAY			DNQEST		KRH

Italicized residues indicate less preferred or “tolerated” residues.

SF 182324.v1

bioRxiv preprint doi: <https://doi.org/10.1101/182324>; this version posted July 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Table IV. HLA Class I Standard Peptide Binding Affinity.

ALLELE	STANDARD PEPTIDE	SEQUENCE	STANDARD BINDING AFFINITY (nM)
A*0101	944.02	YLEPAIAKY	25
A*0201	941.01	FLPSDYFPSV	5.0
A*0202	941.01	FLPSDYFPSV	4.3
A*0203	941.01	FLPSDYFPSV	10
A*0206	941.01	FLPSDYFPSV	3.7
A*0207	941.01	FLPSDYFPSV	23
A*6802	1141.02	FTQAGYPAL	40
A*0301	941.12	KVFPYALINK	11
A*1101	940.06	AVDLYHFLK	6.0
A*3101	941.12	KVFPYALINK	18
A*3301	1083.02	STLPETYVRR	29
A*6801	941.12	KVFPYALINK	8.0
A*2401	979.02	AYIDNYNKF	12
B*0702	1075.23	APRTLVL	5.5
B*3501	1021.05	FPFKYAAAF	7.2
B51	1021.05	FPFKYAAAF	5.5
B*5301	1021.05	FPFKYAAAF	9.3
B*5401	1021.05	FPFKYAAAF	10

Table V. HLA Class II Standard Peptide Binding Affinity.

Allele	Nomenclature	Standard Peptide	Sequence	Binding Affinity (nM)
DRB1*0101	DR1	515.01	PKYVKQNTLKLAT	5.0
DRB1*0301	DR3	829.02	YKTIAFDEEARR	300
DRB1*0401	DR4w4	515.01	PKYVKQNTLKLAT	45
DRB1*0404	DR4w14	717.01	YARFQSQTTLKQKT	50
DRB1*0405	DR4w15	717.01	YARFQSQTTLKQKT	38
DRB1*0701	DR7	553.01	QYIKANSKFIGITE	25
DRB1*0802	DR8w2	553.01	QYIKANSKFIGITE	49
DRB1*0803	DR8w3	553.01	QYIKANSKFIGITE	1600
DRB1*0901	DR9	553.01	QYIKANSKFIGITE	75
DRB1*1101	DR5w11	553.01	QYIKANSKFIGITE	20
DRB1*1201	DR5w12	1200.05	EALIHQLKINPYVLS	298
DRB1*1302	DR6w19	650.22	QYIKANAKFIGITE	3.5
DRB1*1501	DR2w2 $\beta$ 1	507.02	GRTQDENPVVHFFKNIV TPRTPPP	9.1
DRB3*0101	DR52a	511	NGQIGNDPNRDIL	470
DRB4*0101	DRw53	717.01	YARFQSQTTLKQKT	58
DRB5*0101	DR2w2 $\beta$ 2	553.01	QYIKANSKFIGITE	20

The "Nomenclature" column lists the allelic designations used in Tables XIX and XX.

Table VI

HLA-supertype	Allele-specific HLA-supertype members	
	Verified <sup>a</sup>	Predicted <sup>b</sup>
A1	A*0101, A*2501, A*2601, A*2602, A*3201	A*0102, A*2604, A*3601, A*4301, A*8001
A2	A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, A*6901	A*0208, A*0210, A*0211, A*0212, A*0213
A3	A*0301, A*1101, A*3101, A*3301, A*6801	A*0302, A*1102, A*2603, A*3302, A*3303, A*3401, A*3402, A*6601, A*6602, A*7401
A24	A*2301, A*2402, A*3001	A*2403, A*2404, A*3002, A*3003
B7	B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3504, B*3505, B*3506, B*3507, B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, B*7801	B*1511, B*4201, B*5901
B27	B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*3801, B*3901, B*3902, B*7301	B*2701, B*2707, B*2708, B*3802, B*3903, B*3904, B*3905, B*4801, B*4802, B*1510, B*1518, B*1503
B44	B*1801, B*1802, B*3701, B*4402, B*4403, B*4404, B*4001, B*4002, B*4006	B*4101, B*4501, B*4701, B*4901, B*5001
B58	B*5701, B*5702, B*5801, B*5802, B*1516, B*1517	
B62	B*1501, B*1502, B*1513, B*5201	B*1301, B*1302, B*1504, B*1505, B*1506, B*1507, B*1515, B*1520, B*1521, B*1512, B*1514, B*1519

a. Verified alleles includes alleles whose specificity has been determined by pool sequencing analysis, peptide binding assays, or by analysis of the sequences of CTL epitopes.

b. Predicted alleles are alleles whose specificity is predicted on the basis of B and F pocket structure to overlap with the supertype specificity.

TABLE VII

## HBV A01 SUPER MOTIF(With binding information)

Conservancy	Freq.	Protein	Position	Sequence	String	A*0101
95	19	POL	521	AICSWRRAF	XIXXXXXXXF	
95	19	NUC	54	ALRQAILCW	XLXXXXXXW	
80	16	BN	108	AMQWNSTTF	XMXXXXXXF	
100	20	POL	166	ASFCGSPY	XSXXXXXXY	
100	20	POL	166	ASFCGSPYSW	XSXXXXXXXW	
90	18	NUC	19	ASKLCLGW	XSXXXXXXW	
85	17	NUC	19	ASKLCLGWLW	XSXXXXXXXW	
80	16	POL	822	ASPLHVAW	XSXXXXXXW	
100	20	BN	312	CIPISSW	XIXXXXXXXF	
100	20	BN	312	CIPISSWAF	XIXXXXXXXF	
95	19	BN	253	CLFLVLLDY	XLXXXXXXXXY	
95	19	BN	239	CLRRFIIF	XLXXXXXXF	
75	15	BN	239	CLRRFIILF	XLXXXXXXF	
95	19	POL	523	CSVRRAF	XSXXXXXXF	
100	20	BN	310	CTCIPISW	XTXXXXXXXW	
90	18	NUC	31	DIDPYKEF	XIXXXXXXXF	
85	17	NUC	29	DLLDTASALY	XLXXXXXXXXY	11.1000
95	19	BN	196	DSWWTSUNF	XSXXXXXXF	
95	19	NUC	43	ELLSFLPSDF	XLXXXXXXF	
95	19	NUC	43	ELLSFLPSDF	XLXXXXXXF	
95	19	POL	374	ESRLVDF	XSXXXXXXF	
95	19	POL	374	ESRLVDFSF	XSXXXXXXXF	
80	16	BN	248	FILLCLIF	XIXXXXXXXF	
80	16	BN	248	FLFILLCLIF	XLXXXXXXXF	
95	19	BN	258	FLVLLDY	XLXXXXXXY	
95	19	POL	658	FSPTYKAF	XSXXXXXXF	
90	18	X	63	FSSAGPCALRF	XSXXXXXXXF	
100	20	BN	333	FSWLILLVPF	XSXXXXXXXF	
95	19	POL	656	FTFSPTYKAF	XTXXXXXXXF	
95	19	BN	348	FVGLSPTVW	XVXXXXXXXW	
95	19	POL	627	GLLGFAAPF	XLXXXXXXF	
95	19	POL	509	GLSPFLAQF	XLXXXXXXF	
85	17	NUC	29	GMDIDPYKEF	XMXXXXXXXF	0.0017
95	19	NUC	123	GVWIRTPPAY	XVXXXXXXXY	
75	15	POL	569	HLNPNKTKRW	XLXXXXXXXW	
80	16	POL	491	HLYSHPIILGF	XLXXXXXXXF	
85	17	POL	715	HTAELLAACF	XTXXXXXXXF	
95	19	NUC	52	HTALRQAILCW	XTXXXXXXXW	0.0300
100	20	POL	149	HTLWKAGILY	XTXXXXXXXY	
100	20	BN	249	ILLCLIF	XLXXXXXXF	0.0017
80	16	POL	760	ILRGTSFVY	XLXXXXXXY	
90	18	BN	188	ILTIQSLDSW	XLXXXXXXXW	
90	18	POL	625	IVGLGFAAPF	XVXXXXXXXF	
80	16	POL	503	KIPMGVGLSPF	XIXXXXXXXF	
85	17	NUC	21	KLCLGWLW	XLXXXXXXW	
75	15	POL	108	KLIMPARF	XLXXXXXXF	0.0017
75	15	POL	108	KLIMPARFY	XLXXXXXXY	
80	16	POL	610	KLPVNRPIOW	XLXXXXXXXW	
85	17	POL	574	KTKRWGYSUNF	XTXXXXXXXF	0.0680
95	19	POL	55	KVGNFTGLY	XVXXXXXXY	0.0084
95	19	BN	254	LIPLVLLDY	XIXXXXXXXY	
100	20	POL	109	LIMPARFY	XIXXXXXXXY	25.0000
85	17	NUC	30	LIDTASALY	XLXXXXXXY	
80	16	POL	752	LLGCAANW	XLXXXXXXW	
95	19	POL	628	LLGFAAPF	XLXXXXXXF	
100	20	BN	378	LLPIFFCLW	XLXXXXXXXW	
100	20	BN	378	LLPIFFCLWVY	XLXXXXXXXY	
95	19	NUC	44	LLSFLPSDF	XLXXXXXXF	
95	19	NUC	44	LLSFLPSDF	XLXXXXXXXF	
90	18	POL	407	LLSSNLSW	XLXXXXXXW	
95	19	BN	175	LLVLOAGF	XLXXXXXXF	
95	19	BN	175	LLVLOAGFF	XLXXXXXXXF	
100	20	BN	338	LLVPFVQW	XLXXXXXXW	
100	20	BN	338	LLVPFVQWF	XLXXXXXXF	
85	17	NUC	100	LLWFHISCLTF	XLXXXXXXXF	
95	19	NUC	45	LSFLPSDF	XSXXXXXXF	
95	19	NUC	45	LSFLPSDF	XSXXXXXXXF	
95	19	POL	415	LSLDVSAAF	XSXXXXXXXF	4.2000
95	19	POL	415	LSLDVSAAFY	XSXXXXXXXY	
100	20	BN	336	LSLLVPFVQW	XSXXXXXXXW	
100	20	BN	336	LSLLVPFVQWF	XSXXXXXXXF	
95	19	X	53	LSLRGLPVCAF	XSXXXXXXXF	

## HBV A01 SUPER MOTIF(With binding information)

A\*0101

Conservancy	Freq.	Protein	Position	Sequence	String	
95	19	POL	510	LSPFLAQF	XSXXXXXXF	
75	15	BNV	349	LSPTVWLSVIW	XSXXXXXXXXXW	
85	17	POL	742	LSRKYSF	XSXXXXXXF	
85	17	POL	742	LSRKYSFPW	XSXXXXXXXXXW	
75	15	BNV	16	LSVPNPLGF	XSXXXXXXF	
75	15	NUC	137	LTFGRETVLEY	XTXXXXXXXXXY	
90	18	BNV	189	LTIQSLDSW	XTXXXXXXXXXW	
90	18	BNV	189	LTIQSLDSWW	XTXXXXXXXXXW	
90	18	POL	404	LTNLLSSNLSW	XTXXXXXXXXXW	
95	19	BNV	176	LVLQAGFF	XVXXXXXXF	
100	20	BNV	339	LVPFVQWF	XVXXXXXXF	
100	20	POL	377	LWDFSQF	XVXXXXXXF	0.0810
85	17	BNV	360	MMWYWGPSLY	XMXXXXXXXXXY	0.8500
75	15	X	103	MSTTDLEAY	XSXXXXXXXXY	
75	15	X	103	MSTTDLEAYF	XSXXXXXXXXF	
95	19	POL	42	NLGNLNVSIW	XLXXXXXXXXXW	
90	18	POL	406	NLLSSNLSW	XLXXXXXXXXXW	
95	19	POL	45	NLNVSIW	XLXXXXXXXXXW	
75	15	BNV	15	NLSVPNPLGF	XLXXXXXXXXF	0.0005
90	18	POL	738	NSVLSRKY	XSXXXXXXXXY	0.0078
100	20	BNV	380	PIFFCLWVY	XIXXXXXXY	
100	20	BNV	314	PIPSSWAF	XIXXXXXF	0.0190
100	20	POL	124	PLDKGIKPY	XLXXXXXXXXY	0.1600
100	20	POL	124	PLDKGIKPY	XLXXXXXXXXY	
100	20	BNV	377	PLLPIFFCLW	XLXXXXXXXXXW	
95	19	BNV	174	PLLVLQAGF	XLXXXXXXXXF	
95	19	BNV	174	PLLVLQAGFF	XLXXXXXXXXF	
80	16	POL	505	PMGVGLSPF	XMXXXXXXF	0.7700
85	17	POL	797	PTTGRTSLY	XTXXXXXXXXY	
75	15	BNV	351	PTVWLSVIW	XTXXXXXXXXXW	
85	17	POL	612	PVNRPIOW	XVXXXXXXW	
95	19	POL	685	QVFADATPTG	XVXXXXXXXXXW	
90	18	POL	624	RIVGLGF	XIXXXXXF	
75	15	POL	106	RLKLIMPARF	XLXXXXXXXXXW	
75	15	POL	106	RLKLIMPARFY	XLXXXXXXXXXY	
95	19	POL	376	RLWDFSQF	XLXXXXXXXXF	
90	18	POL	353	RTPARVTGGVF	XTXXXXXXXXXW	
100	20	POL	49	SIPWTHKVGNF	XIXXXXXXXXXF	
95	19	BNV	194	SLDSWWTSLNF	XLXXXXXXXXXW	
95	19	POL	416	SLDVSAAF	XLXXXXXW	17.2000
95	19	POL	416	SLDVSAAFY	XLXXXXXXXXY	
100	20	BNV	337	SLLVPPVQWF	XLXXXXXXXXXW	
100	20	BNV	337	SLLVPPVQWF	XLXXXXXXXXXW	
95	19	X	54	SLRGLPVCFA	XLXXXXXXXXF	
90	18	X	64	SSAGPCALRF	XSXXXXXXXXXW	
75	15	X	104	STTDLEAY	XTXXXXXY	
75	15	X	104	STTDLEAYF	XTXXXXXW	
75	15	BNV	17	SVNPLGF	XVXXXXXXF	
90	18	POL	739	SVVLSRKY	XVXXXXXY	
85	17	POL	739	SVVLSRKYSF	XVXXXXXXXXXW	
90	18	BNV	190	TIPQSLDSW	XIXXXXXXW	
90	18	BNV	190	TIPQSLDSWW	XIXXXXXXXXXW	
100	20	POL	150	TLWKAGILY	XLXXXXXXXXY	0.0017
75	15	X	105	TTDLEAYF	XTXXXXXW	
85	17	POL	798	TTGRTSLY	XTXXXXXY	
80	16	NUC	16	TVQASKLCLGW	XVXXXXXXXXXW	
75	15	BNV	352	TVWLSVIW	XVXXXXXXW	
85	17	POL	741	VLSRKYSF	XLXXXXXXF	
85	17	POL	741	VLSRKYSFPW	XLXXXXXXXXXW	
85	17	POL	740	VVLSRKYSF	XVXXXXXXXXXW	
80	16	POL	759	WILRGTSF	XIXXXXXF	0.0023
80	16	POL	759	WILRGTSFVY	XIXXXXXXXXXY	
95	19	NUC	125	WIRTPPAY	XIXXXXXY	
80	16	POL	751	WLLGCAANW	XLXXXXXXW	
95	19	POL	414	WLSLOVSAAF	XLXXXXXXXXXW	
95	19	POL	414	WLSLDVSAAFY	XLXXXXXXXXXY	
100	20	BNV	335	WLSLLVPF	XLXXXXXW	
100	20	BNV	335	WLSLLVPPVQWF	XLXXXXXXXXXW	0.0810
85	17	NUC	26	WLWGMIDPY	XLXXXXXXXXXY	
95	19	BNV	237	WMCLRRFIIF	XMXXXXXXF	
85	17	BNV	359	WMMWYWGFS	XMXXXXXXXXXY	
100	20	POL	52	WTHKVGNF	XTXXXXXW	

# HBV A01 SUPER MOTIF(With binding information)

Conservancy	Freq.	Protein	Position	Sequence	String	A*0101
100	20	POL	122	YLPLDKGIKPY	XLXXXXXXXXY	
90	18	NUC	118	YLVSGVW	XLXXXXXW	
80	16	POL	493	YSHPIILGF	XSXXXXXXF	
85	17	POL	580	YSLNFMGY	XSXXXXXY	
			148			

Table VIII

HBV A02 Super Motif with Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0201	A*0202	A*0203	A*0206	A*6802
85	17	POL	721	AACFARSRSGA	11					
85	17	POL	431	AAMPILIV	8					
80	16	POL	756	AANWILRG	9					
95	19	POL	632	AAPETOCGYPA	11					
95	19	POL	521	AICSVVRA	9	0.0001				
90	18	NJC	58	AICWGEL	8					
90	18	NJC	58	AICWGEL	9					
95	19	POL	642	ALMPLYACI	9	0.5000	0.0340	3.3000	0.0250	0.0470
80	16	BNV	108	AMOWNSIT	8					
75	15	X	102	AMSTIDLEA	9	0.0013				
95	19	POL	516	AOFTSAICSV	10					
95	19	POL	516	AOFTSAICSV	11					
95	19	POL	690	ATPTGWGL	8					
80	16	POL	690	ATPTGWGL	9					
75	15	POL	690	ATPTGWGL	10					
95	19	POL	397	AVPNLOSL	8					
95	19	POL	397	AVPNLOSL	9	0.0001				
95	19	POL	755	CAANWILRG	11					
80	16	POL	61	CAFSAGPCA	10					
95	19	X	61	CAFSAGPCA	10	0.0001				
95	19	X	61	CAFSAGPCA	11					
90	18	X	69	CALRTSA	8					
100	20	BNV	312	CIPPSWA	9	0.0010				
80	16	BNV	312	CIPPSWAFA	11					
90	18	POL	533	CLAFSYMDV	10	0.0008				
90	18	POL	533	CLAFSYMDV	11					
85	17	NJC	23	CLGWLWGM	8					
85	17	NJC	23	CLGWLWGM	10	0.0093				
100	20	BNV	253	CLFLVL	8	0.0002				
100	20	BNV	253	CLFLVL	9	0.0006				
95	19	BNV	239	CLRRHFL	9	0.0002				
75	15	BNV	239	CLRRHFL	11	0.0004				
90	18	NJC	107	CLTGHET	8					
90	18	NJC	107	CLTGHET	9	0.0001				
80	16	X	7	COLDPARDV	9					
80	16	X	7	COLDPARDV	10					
85	17	POL	622	COBWGL	8					
85	17	POL	622	COBWGL	11					
95	19	POL	684	COVFADAT	8					
95	19	POL	684	COVFADAT	10					
100	20	BNV	310	CTCIPPSWA	11	0.0001				
95	19	POL	689	DATPTGWGL	9					
80	16	POL	689	DATPTGWGL	10					
75	15	POL	689	DATPTGWGL	11					
90	18	NJC	31	DIDPYKEGA	10					
85	17	NJC	29	DLDYASA	8					

# IBV A02 Super Motif with Binding Information

Conservancy	Freq	Protein	Position	Sequence	AA	A*0201	A*0202	A*0203	A*0206	A*6802
85	17	N/C	29	DLIDTASAL	9	0.0001				
95	19	POL	40	DNLGNLW	9	0.0004				
95	19	POL	40	DNLGNLWSI	11					
80	16	N/C	32	DTASALYREA	10					
80	16	N/C	32	DTASALYREAL	11					
95	19	X	14	DWLQAPV	8					
95	19	X	14	DWLQAPVGA	10	0.0001				
90	18	POL	541	DVLGAKSV	9	0.0003				
100	20	POL	17	EAGPLEEL	9	0.0001				
80	16	X	122	ELGEERL	8					
90	18	POL	122	ELLAACFA	8					
75	15	N/C	142	ETVLEYLV	8					
95	19	POL	687	FADATPTGWGL	11					
85	17	POL	724	FARSRGA	8					
80	16	POL	821	FASPLHYA	8					
95	19	POL	396	FAVPLIOL	9					
95	19	POL	396	FAVPLIOLST	10	0.0003				
80	16	EW	243	FIFLFL	8	0.0006				
80	16	EW	243	FIFLFL	9	0.0002				
80	16	EW	243	FIFLFL	10	0.0012				
80	16	EW	248	FILLCL	8	0.0003				
80	16	EW	248	FILLCLFL	10	0.0280				
80	16	EW	246	FILLCLFL	11	0.0010				
80	16	EW	246	FLILLCL	9	0.0002				
80	16	EW	246	FLILLCL	10	0.0013				
75	15	EW	171	FLGPLVL	8					
75	15	EW	171	FLGPLVLOA	10	0.0190				
95	19	POL	513	FLAOFISA	9	0.2400				
95	19	POL	513	FLAOFISA	10	0.2100				
95	19	POL	562	FLSLGHL	9	0.6500				
80	16	EW	183	FLTRLIT	8					
80	16	EW	183	FLTRLIT	9	0.5100				
95	19	EW	256	FLVLDYOGM	11		0.0430	8.0000	0.2000	0.0010
100	20	POL	363	FLYDKNPNT	10	0.0012				
95	19	POL	656	FTFSPTYKA	9	0.0056	0.0150	0.0031	0.0008	7.3000
95	19	POL	656	FTFSPTYKAF	11					
95	19	POL	59	FTGLYSST	8					
90	18	POL	59	FTGLYSSTV	9	0.0005				
95	19	POL	635	FTOQGYPA	8					
95	19	POL	635	FTOQGYPAL	9	0.0009				
95	19	POL	635	FTOQGYPALM	10	0.0024				
95	19	POL	518	FTSAICSV	8					
95	19	POL	518	FTSAICSVV	9	0.0090				
95	19	EW	346	FVGLSPTV	8					
95	19	EW	346	FVGLSPTVWL	10	0.0008				
90	18	X	132	FVGLGCRHKL	10	0.0030				



# HBV A02 Super Motif with Binding Information

Conservancy	Freq	Protein	Position	Sequence	AA	A*0201	A*0202	A*0203	A*0206	A*6802
90	18	X	132	FMGGCRHLV	11					
95	19	BN	342	PCQMPVGL	8					
95	19	BN	342	PCQMPVGLSPT	11					
90	18	POL	766	FVYVPSAL	8					
90	18	POL	766	FVYVPSALNPA	11					
95	19	X	50	GAHLSRLGL	9	0.0001				
90	18	X	50	GAHLSRLGLPV	11					
85	17	POL	545	GAKSVQHL	8					
85	17	POL	545	GAKSVQHLESL	11					
75	15	POL	567	GHLNPNKT	9					
90	18	POL	155	GILYKRET	8					
90	18	POL	155	GILYKRETT	9					
85	17	POL	682	GLOOVFADA	9	0.0024				
85	17	POL	682	GLOOVFADAT	10					
95	19	POL	627	GILGFAPFT	10	0.0049				
85	17	BN	62	GILGWSFOA	9	0.4000				
95	19	X	57	GLPVCAFSSA	10	0.0008				
95	19	POL	509	GLSPFLA	8					
95	19	POL	509	GLSPFLAQT	11					
100	20	BN	348	GLSPTVWL	8	0.0036				
75	15	BN	348	GLSPTVWLSV	10	0.2800				
75	15	BN	348	GLSPTVWLSVI	11	0.0036				
90	18	BN	265	GMLPVCPPL	8		0.0003			
90	18	POL	735	GTDSNVL	8					
75	15	BN	13	GTNLSVNPPL	10					
80	16	POL	763	GTSFVVPSPA	10					
80	16	POL	763	GTSFVVPSPAL	11					
80	16	POL	507	GVGLSPFL	8					
80	16	POL	507	GVGLSPFLA	9	0.0002				
80	16	POL	507	GVGLSPFLA	10					
95	19	NJC	123	GVWIRTPPA	9	0.0030				
90	18	NJC	104	HISCLTFGRET	11					
80	16	POL	435	HLVGSGL	9	0.0031				
90	18	X	52	HLSLRGLPV	9	0.0014				
90	18	X	52	HLSLRGLPVCA	11					
80	16	POL	491	HLYSHIP	8					
80	16	POL	491	HLYSHIPIL	9	0.2200				
85	17	POL	715	HTAELLA	8		0.0003			
85	17	POL	715	HTAELLAACFA	11			0.9300	0.0017	0.0530
100	20	NJC	52	HTALROAI	8					
95	19	NJC	52	HTALROAIL	9	0.0001				
100	20	POL	149	HTLWKAGI	8	0.0001				
100	20	POL	149	HTLWKAGIL	9	0.0004				
80	16	BN	244	IIFLFIILL	8	0.0002				
80	16	BN	244	IIFLFIILL	9	0.0002				
80	16	BN	244	IIFLFIILLCT	11					

# HBV A02 Super Motif with Binding Information

Conservancy	Freq	Protein	Position	Sequence	AA	A*0201	A*0202	A*0203	A*0206	A*6802
80	16	POL	497	ILGFRKI	8					
80	16	POL	497	ILGFRKIPM	10					
90	18	NLC	59	ILCWGEIM	8					
80	16	POL	498	ILGFRKIPM	9	0.0002				
100	20	BW	249	ILLCLIFL	9	0.0015				
100	20	BW	249	ILLCLIFLL	10	0.0190	0.0001	0.0002	0.0013	0.0015
100	20	BW	249	ILLCLIFLV	11	0.0056				
80	16	POL	760	ILRGTSFV	8					
80	16	POL	760	ILRGTSFVYV	10	0.0160				
100	20	NLC	139	ILSTLPET	8					
100	20	NLC	139	ILSTLPETT	9	0.0001				
100	20	NLC	139	ILSTLPETTV	10	0.0210	0.0085	0.0770	0.0031	0.0067
100	20	NLC	139	ILSTLPETTVV	11					
95	19	BW	188	ILTIPOS	8					
90	18	POL	156	ILYKRETT	8					
90	18	POL	625	IVGLGFA	8					
90	18	POL	625	IVGLGFAA	9	0.0009				
90	18	POL	153	KAGILYKRETT	10					
90	18	POL	153	KAGILYKRETT	11					
80	16	POL	503	KIPMGVGL	8					
85	17	NLC	21	KLCGLMWLGM	10	0.0001				
95	19	POL	489	KLHLYSHP	9	0.0690	0.0340	2.7000	0.0059	0.0015
80	16	POL	489	KLHLYSHPIL	10					
80	16	POL	610	KLHLYSHPIL	11					
80	16	POL	610	KLPVNRPI	8					
95	19	POL	653	KOAFISPT	9					
95	19	POL	574	KTKRWGYSL	9	0.0001				
85	17	POL	620	KVCOHIVGL	9	0.0003				
85	17	POL	620	KVCOHIVGL	10	0.0001				
95	19	POL	55	KVGNFTGL	8					
85	17	X	91	KVLHKRTL	8					
85	17	X	91	KVLHKRTGL	10	0.0004				
90	18	POL	534	LAFSYMDV	9	0.0002				
90	18	POL	534	LAFSYMDV	10	0.0003				
90	18	POL	534	LAFSYMDVVL	11					
95	19	POL	515	LAOFTSAI	8					
95	19	POL	515	LAOFTSAICSV	11					
100	20	BW	254	LIFLLVLL	8	0.0025				
95	19	POL	514	LLAQTSA	8					
95	19	POL	514	LLAQTSAI	9	0.1000	0.2700	3.7000	0.0026	0.7900
100	20	BW	251	LLCLIFLL	8	0.0004				
100	20	BW	251	LLCLIFLV	9	0.0048				
100	20	BW	251	LLCLIFLV	10	0.0075				
100	20	BW	251	LLCLIFLVLL	11	0.0013				
85	17	NLC	30	LLDTASAL	8					
95	19	BW	260	LLDYOGML	8	0.0004				

# HBV A02 Super Motif with Binding Information

Conservancy	Freq	Protein	Position	Sequence	AA	A*0201	A*0202	A*0203	A*0206	A*6802
90	18	EW	260	LLDYOGMLPV	10	0.0980	0.0001	0.0200	0.0067	0.0009
80	16	POL	752	LLGCANWIL	9	0.0011				
80	16	POL	752	LLGCANWIL	10	0.0140				
95	19	POL	628	LLGFAPFT	9	0.0008				
85	17	EW	63	LLGWSPOA	8					
75	15	EW	63	LLGWSPOA	11					
100	20	EW	250	LLLCIFL	8	0.0006				
100	20	EW	250	LLLCIFL	9	0.0065				
100	20	EW	250	LLLCIFLV	10	0.0036				
100	20	EW	250	LLLCIFLV	11	0.0005				
100	20	EW	378	LLPIFCL	8	0.0055				
100	20	EW	378	LLPIFCLWV	10	0.0320	0.0008	0.0150	0.0008	0.0005
95	19	POL	563	LLSLGHL	8					
90	18	POL	407	LLSSNLSQL	9	0.0110	0.0780	3.9000	0.0027	0.0100
90	18	POL	407	LLSSNLSQL	11					
80	16	EW	184	LLTRILTI	8	0.0026				
80	16	POL	436	LLVSSQL	8	0.0050				
95	19	EW	257	LLVLDYOGM	10					
95	19	EW	257	LLVLDYOGM	11	0.0310	0.0037	0.0045	0.0015	0.0110
90	18	EW	175	LLVLDYOGM	10	0.0074				
95	18	EW	175	LLVLDYOGM	11	0.6700	0.3800	1.7000	0.2900	0.1400
90	19	EW	338	LLVLDYOGM	10					
90	18	EW	100	LLVLDYOGM	9	0.0130	0.0002	0.0420	0.0031	0.0098
85	17	NJC	100	LLVLDYOGM	10					
95	19	POL	643	LLVLDYOGM	8					
95	19	EW	178	LLVLDYOGM	8					
95	19	EW	178	LLVLDYOGM	9					
80	16	EW	178	LLVLDYOGM	11					
100	20	POL	401	LLVLDYOGM	8					
95	19	NJC	108	LLVLDYOGM	8					
75	15	NJC	137	LLVLDYOGM	9					
90	18	POL	404	LLVLDYOGM	9					
80	16	EW	185	LLVLDYOGM	11					
85	17	POL	99	LLVLDYOGM	9					
100	20	POL	364	LLVLDYOGM	9	0.0001				
95	19	EW	258	LLVLDYOGM	9	0.0001				
95	19	EW	258	LLVLDYOGM	10	0.0001				
90	18	EW	176	LLVLDYOGM	10	0.0096				
90	18	EW	176	LLVLDYOGM	11	0.0022				
90	18	EW	176	LLVLDYOGM	9					
95	19	EW	339	LLVLDYOGM	9	0.0420	0.0150	0.0048	0.0079	2.8000
95	19	EW	339	LLVLDYOGM	11					
90	18	NJC	119	LLVLDYOGM	8	0.0004				
90	18	NJC	119	LLVLDYOGM	10					
85	17	EW	360	LLVLDYOGM	9					
75	15	NJC	1	LLVLDYOGM	8	0.6400				

# IBY A02 Super Motif with Binding Information

Conservancy	Freq	Protein	Position	Sequence	AA	A*0201	A*0202	A*0203	A*0206	A*6802
100	20	NLC	136	NAPILSTL	8					
100	20	NLC	136	NAPILSTPET	11					
95	19	POL	42	NLGNLWSI	9	0.0047				
90	18	POL	406	NILNSLSWL	10	0.0016				
95	19	POL	45	NLNSIPVT	9	0.0005				
100	20	POL	400	NLOSLNL	8					
100	20	POL	400	NLOSLNLL	9	0.0047				
75	15	BW	15	NLSVPNPL	8					
90	18	POL	411	NLSWLSLDV	9	0.0650				
90	18	POL	411	NLSWLSLDVSA	11		0.0051	0.6400	0.0016	0.0990
100	20	POL	47	NVSPVTHKV	10	0.0001				
100	20	POL	430	PAAMPPLL	8					
85	17	POL	430	PAAMPPLV	9					
90	18	POL	775	PAOPSRGRL	10					
90	18	BW	131	PAGSSSGT	9					
90	18	BW	131	PAGSSSGTV	10					
95	19	POL	641	PALMPLYA	8					
95	19	POL	641	PALMPLVACI	10	0.0001				
75	15	X	145	PAPCNFT	8					
75	15	X	145	PAPCNFTSA	10					
80	16	X	11	PARDMLCL	8					
75	15	X	11	PARDMLCLRPV	11					
90	18	POL	355	PARVTGGV	8					
90	18	POL	355	PARVTGGVFL	10					
90	18	POL	355	PARVTGGVFLV	11					
95	19	NLC	130	PAYRPPNA	8					
95	19	NLC	130	PAYRPPNAPL	10	0.0001				
85	17	POL	616	PIDWKVCOAI	11	0.0001				
85	17	POL	616	PIDWKVCOAIRV	11					
100	20	BW	380	PIFCLWV	8					
100	20	BW	380	PIFCLWYI	10	0.0004				
85	17	POL	713	PIHTAELL	8					
85	17	POL	713	PIHTAELLA	9					
85	17	POL	713	PIHTAELLA	10					
80	16	POL	496	PILGFRKI	9	0.0001				
80	16	POL	496	PILGFRKIPM	11					
100	20	NLC	138	PILSTLPET	9	0.0001				
100	20	NLC	138	PILSTLPETT	10	0.0001				
100	20	NLC	138	PILSTLPETT	11	0.0001				
80	16	BW	314	PIPSWAF	9					
95	19	POL	20	PLEEELPRL	9	0.0003				
90	18	POL	20	PLEEELPRLA	10	0.0001				
95	19	BW	10	PLGFFDQOL	10	0.0002				
100	20	POL	427	PLHPPAAMPPL	10	0.0001				
100	20	POL	427	PLHPPAAMPPL	11					

# HBV A02 Super Motif with Binding Information

Conservancy	Freq	Protein	Position	Sequence	AA	A*0201	A*0202	A*0203	A*0206	A*6802
100	20	BN	377	PLPFFQL	9	0.0650				
100	20	BN	377	PLPFFQLW	11					0.0047
90	18	BN	174	PLVLQAGFL	11	0.0008	0.0001			
80	16	POL	711	PLPHIAEL	9	0.0004				
80	16	POL	711	PLPHIAELL	10	0.0001				
80	16	POL	711	PLPHIAELLA	11					
75	15	POL	2	PLSYCHFRKL	10	0.0001				
75	15	POL	2	PLSYCHFFKLL	11					
85	17	POL	98	PLTVNEKRL	10	0.0001				
80	16	POL	505	PMGVGLSPFL	10	0.0001				
80	16	POL	505	PMGVGLSPFL	11					
95	19	BN	106	POAMOWNST	9					
80	16	BN	106	POAMOWNST	10					
90	18	BN	192	POSLSWWT	9					
90	18	BN	192	POSLSWWTSL	11					
75	15	POL	692	PTGWGLAI	8					
80	16	BN	219	PTSNHPT	8					
85	17	POL	797	PTGRTSL	8					
85	17	POL	797	PTGRTSLYA	10					
80	16	NJC	15	PTVOASKL	8					
80	16	NJC	15	PTVOASKLCL	10					
75	15	BN	351	PTVWLSVI	8					
75	15	BN	351	PTVWLSVIM	10					
95	19	X	59	PVCAFSSA	8					
85	17	POL	612	PVNRPIDWKV	10					
95	19	POL	654	QAFESPT	8	0.0002				
95	19	POL	654	QAFESPTYKA	11					
95	19	BN	179	QAGFLIT	8					
80	16	BN	179	QAGFLITRI	10					
80	16	BN	179	QAGFLITRIL	11					
90	18	NJC	57	QALCWGEL	9					
90	18	NJC	57	QALCWGELM	10					
95	19	BN	107	QAMOWNST	8					
80	16	BN	107	QAMOWNST	9					
80	16	NJC	18	QASKLQMWL	10					
80	16	X	8	QLDPADIV	8	0.0001				
80	16	X	8	QLDPADIVL	9	0.0001				
80	16	X	8	QLDPADIVLQ	11	0.0001				
90	18	NJC	99	QLWFHISCL	10	0.0060				
85	17	NJC	99	QLWFHISCLT	11					
95	19	POL	685	QVFADATPT	9	0.0001				
95	19	POL	528	RAFPICLA	8					
80	16	BN	187	RILTIPOS	9					
90	18	POL	624	RIVGLGFA	9	0.0010				
90	18	POL	624	RIVGLGFAA	10					
75	15	POL	106	RLKLMPA	8					

# HBV A02 Super Motif with Binding Information

Conservancy	Freq	Protein	Position	Sequence	AA	A*0201	A*0202	A*0203	A*0206	A*6802
90	18	NJC	56	FOALCWGEL	10					
90	18	NJC	56	FOALCWGELM	11					
90	18	NJC	98	ROLWTHI	8					
90	18	NJC	98	ROLWTHISCL	11					
85	17	BNV	88	ROSGROPT	8					
90	18	POL	353	RTPARVTGGV	10					
95	19	NJC	127	RTPPAYRPPNA	11					
95	19	POL	36	RVAEDLNL	8					
90	18	POL	36	RVAEDLNLGNL	11					
80	16	POL	818	RVHFAFSL	8					
75	15	POL	818	RVHFAFSLV	10	0.0001				
75	15	POL	818	RVHFAFSLHVA	11					
100	20	POL	357	RVTGGVFL	8					
100	20	POL	357	RVTGGVFLV	9	0.0041				
90	18	X	65	SAGCALRFT	10					
95	19	POL	520	SAICSVVRRA	10	0.0001				
90	18	NJC	35	SALYREAL	8					
100	20	POL	49	SIPWTHKV	8					
95	19	BNV	194	SLDSWMTSL	9	0.0023				
75	15	POL	565	SLGHLNPNKT	11					
95	19	BNV	337	SILVPRVQMFV	11					
75	15	POL	581	SILNFMGV	8					
75	15	POL	581	SILNFMGVY	9	0.0038				
95	19	X	54	SLRGLPYCA	9	0.0007				
90	18	POL	403	SLTNLLSNL	10	0.0014				
75	15	BNV	216	SCSPISNHSPT	11					
75	15	BNV	280	STGPKCT	9					
100	20	NJC	141	STLPETTV	8					
100	20	NJC	141	STLPETTV	9	0.0019				
80	16	BNV	85	STNFOGFOPT	11					
85	17	POL	548	SVCHLESL	8					
80	16	BNV	330	SVRFWSLSL	9	0.0001				
80	16	BNV	330	SVRFWSLSL	10	0.0004				
80	16	BNV	330	SVRFWSLSL	11					
90	18	POL	739	SVVLSRKYT	9					
95	19	POL	524	SVVRRAPPHQL	11					
85	17	POL	716	TAELLAACFA	10					
95	19	NJC	53	TALROAIL	8					
80	16	NJC	33	TASALYREA	9					
80	16	NJC	33	TASALYREAL	10					
90	18	BNV	190	TIPOSIDSWNT	11					
100	20	NJC	142	TLPETTV	8					
100	20	POL	150	TLWKAGIL	8					
95	19	POL	636	TOCGYPAL	8					
95	19	POL	636	TOCGYPALM	9					
95	19	POL	636	TOCGYPALMPL	11					

HBV A02 Super Motif with Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0201	A*0202	A*0203	A*0206	A*6802
85	17	POL	798	TTGRTSLYA	9					
75	15	BN	278	TTSTGPCKT	9					
75	15	BN	278	TTSTGPCKTCT	11					
85	17	POL	100	TNNEKRL	8					
80	16	MUC	16	TYQASKLCL	9	0.0002				
75	15	BN	352	TVWLSVWMM	9	0.0002				
95	19	POL	37	VAEDNLGNL	10	0.0001				
95	19	X	15	VLCLRPVGA	9	0.0014				
85	17	POL	543	VLGAKSVCHL	10	0.0001				
90	18	X	133	VLGGCHKL	9	0.0009				
90	18	X	133	VLGGCHKL	10	0.0001				
85	17	X	92	VLHKRTLGL	9	0.0012				
95	19	BN	259	VLDYOGM	8					
95	19	BN	259	VLDYOGML	9	0.0440	0.0001	0.0210	0.0009	0.0002
90	18	BN	259	VLDYOGMLPV	11	0.5800	0.2200	4.9000	0.3400	0.0170
95	19	BN	177	VLOAGFL	8	0.0019				
95	19	BN	177	VLOAGFLL	9	0.0660				
95	19	BN	177	VLOAGFLL	10	0.0011				
80	16	MUC	17	VOASKLCL	8					
80	16	MUC	17	VOASKLCLGWL	11					
95	19	BN	343	VQMFVGLSPT	10					
95	19	BN	343	VQMFVGLSPTV	11					
100	20	POL	358	VTGVRV	8					
90	18	POL	542	VVLGAKSV	8					
80	16	POL	542	VVLGAKSVCHL	11					
90	18	POL	740	VVLSRKYT	8					
95	19	POL	525	VVRAPFHCL	10	0.0003				
95	19	POL	525	VVRAPFHCLA	11					
80	16	POL	759	WILRGTSFV	9	0.0270				
80	16	POL	759	WILRGTSFV	11					
80	16	POL	751	WLLGCAANWI	10	0.0053				
80	16	POL	751	WLLGCAANWIL	11					
100	20	POL	414	WLSLDVSA	8					
95	19	POL	414	WLSLDVSA	9	0.0059				
100	20	BN	335	WLSLVPFV	9	1.1000	0.0380	7.2000	0.3600	0.0310
95	19	BN	237	WMCLRRRI	8					
95	19	BN	237	WMCLRRRI	9	0.0005				
95	19	BN	237	WMCLRRRI	11	0.0019				
85	17	BN	359	WMWMMWGPFL	10	0.0009				
100	20	BN	52	WTHKVGNT	9	0.0001				
95	19	POL	52	WTHKVGNTGL	11					
100	20	POL	147	YLHTLWKA	8					
100	20	POL	147	YLHTLWKA	10	0.0160	0.0005	0.5600	0.0010	0.0320
100	20	POL	147	YLHTLWKA	11					
100	20	POL	122	YLPDQGI	8					
90	18	MUC	118	YLVSGWMI	9	0.3800				

HBV A02 Super Motif with Binding Information

# HBV A02 Super Motif with Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0201	A*0202	A*0203	A*0206	A*6802
90	18	NLC	118	YLVSEGWIRT	11					
90	18	POL	538	YMDDVVLGA	9	*	0.0250	0.0001	0.0024	0.0001
90	18	ENV	263	YQGMPLVCP	10					0.0002
75	15	POL	5	YQHFRKLL	8					
75	15	POL	5	YQHFRKLL	9			1		
75	15	POL	5	YQHFRKLL	10					
85	17	POL	746	YTSFPWLL	8					
75	15	POL	746	YTSFPWLLGCA	11					
90	18	POL	768	YVPSALNPA	9			0.0039		
				423	3	45				



HBV A03 SUPER MOTIF (with blinding information)

Conservancy	Frequency	Protein	Position	Sequence	P2	C-term	AA	A*0301	A*1101	A*3101	A*3301	A*6801
85	17	POL	721	AACFARSR	A	R	8	0.0004	0.0003	0.0056	0.0035	0.0014
95	19	POL	521	AICSVVIR	I	R	8	-0.0002	0.0003	0.0014	-0.0009	0.0006
90	18	POL	772	ALNPADPSR	L	R	10	0.0003	0.0001	0.0009	0.0001	0.0004
85	17	X	70	ALRFTSAR	L	R	8	0.0047	0.0009	0.0450	0.0230	0.0004
80	16	POL	822	ASPLVAVIR	S	R	9					
75	15	ENV	84	ASTNRQSR	S	R	9	0.0009	0.0002	0.0088	0.0008	0.0001
80	18	POL	755	CAANWILR	A	R	8					
85	17	X	69	CALRITSAR	A	R	9	0.0034	0.0230	1.5000	8.0000	0.7300
90	18	X	17	CLRPVAGESR	L	R	10	0.0011	0.0001	0.0520	0.0250	0.0440
100	20	NJC	48	CSRHITALR	S	R	9	0.0028	0.0001	0.0003	0.0003	0.0410
85	17	NJC	29	DILDTASALYR	L	R	11	0.0042	-0.0003	-0.0012	3.7000	0.0410
85	17	NJC	32	DTASALYR	T	R	8	0.0004	-0.0002	-0.0009	0.0018	0.0009
95	19	POL	17	EAGRLEELPR	A	R	11	-0.0009	-0.0003	-0.0012	0.0015	0.0110
90	18	POL	718	ELAACFAR	L	R	8	0.0002	0.0004	0.0004		
85	17	POL	718	ELAACFARSR	L	R	11	0.0062	0.0016	0.0200	0.2000	0.1600
95	19	NJC	174	ETTVERRR	T	R	8	0.0003	-0.0002	-0.0009	0.1400	0.0027
80	16	NJC	174	ETTVERRRGR	T	R	10	0.0003	0.0001			
80	16	POL	821	FASPLVAVIR	A	R	10					
90	18	X	83	FSSAGPCALR	S	R	10					
95	19	POL	856	FTFSRPTK	T	R	8	0.0160	0.0100	0.0023	0.2100	0.0590
95	18	POL	518	FTSACSVIR	T	R	10	0.0003	0.0003			
85	18	POL	518	FTSACSVIR	T	R	11	0.0055	0.0082	0.0170	0.0350	1.5000
90	15	X	132	FLGSGCTHK	V	K	8	0.0430	0.0090			
75	15	POL	567	GHLPNPK	I	K	10	0.0025	0.0011	0.0009	0.0009	0.0003
75	15	POL	567	GHLPNPKTK	I	K	11					
75	15	POL	567	GHLPNPKTKR	I	K	11					
85	17	NJC	29	GKMDPYK	M	R	8	0.0006	0.0004	-0.0009	-0.0009	0.0001
90	18	POL	736	GTDSWLSR	T	R	10	0.0010	0.0420	0.0030	0.0019	0.0008
90	18	POL	735	GTDSWLSHR	T	R	11	0.0140	0.5600	-0.0002	-0.0005	0.0001
95	19	NJC	123	GVWHTTPAYR	V	R	11	0.1900	0.1700	6.8000	0.7300	0.6600
90	18	NJC	104	HISLITFGR	I	R	9	0.0160	0.0065			
75	15	POL	568	HLNPKTK	L	R	8					
75	15	POL	568	HLNPKTKR	L	R	8	0.0025	0.0001			
100	20	NJC	148	HTLWKAGILYK	T	K	11	0.0370	0.4400	0.0370	0.0720	0.1900
90	18	NJC	105	ISCLITGR	S	R	8	0.0004	0.0002	0.0017	-0.0009	0.0017
100	20	POL	153	KAGILYK	A	R	8	0.0002	-0.0002	0.0015	-0.0009	0.0001
80	16	POL	810	KLPVNRIDPK	L	K	11					
75	15	X	130	KVFLGGR	V	R	9	0.0420	0.0820	0.5000	0.0710	0.0030
85	17	POL	720	LAACFARSR	A	R	9	0.0058	0.0065			
90	18	POL	718	LLAACFAR	L	R	8	0.0024	0.0003	0.0015	0.0028	0.0064
85	17	POL	719	LLAACFARSR	L	R	10					
85	17	NJC	30	LLDTASALYR	L	R	10	0.0050	0.0002			
60	16	POL	752	LLGCAANWILR	L	R	11					
75	15	POL	564	LSGLHLPNPK	S	K	11					
95	18	NJC	169	LSLTPETTVIR	S	R	11	-0.0008	0.0008	-0.0012	-0.0023	0.0078
75	15	POL	3	LSYCHRRK	S	K	8					
85	17	POL	88	LYNDRGR	T	R	8	-0.0002	-0.0002	-0.0009	-0.0009	0.0001
90	18	NJC	118	LYSFGWMR	V	R	9	0.0028	0.0120			
100	20	POL	377	LYWDFSCFSR	V	R	10	0.0016	0.3600	0.0260	0.2300	0.4900
75	15	X	103	MSITDLEAYFK	S	K	11					
90	18	NJC	75	NEDPASR	L	R	8	-0.0002	-0.0002	-0.0009	-0.0009	0.0001

# HBV A03 SUPER MOTIF (With binding Information)

Conservancy	Frequency	Protein	Position	Sequence	P2	C-term	AA	A*0301	A*1101	A*3101	A*3301	A*0801
85	19	POL	45	NLVNIPWTHK	L	K	11	-0.0009	0.0005	-0.0012	-0.0023	0.0019
90	18	POL	738	NSVMSRK	S	K	B	0.0006	0.0010	-0.0008	-0.0009	0.0007
100	20	POL	47	NSIPWTHK	V	K	B	0.0020	0.0570	0.0100	0.0002	0.0320
80	18	POL	775	PADPSRGR	A	R	9	0.0008	0.0002	0.0004	0.0015	0.0002
80	16	X	11	PADVLCR	A	R	9	0.0002	0.0002	0.0100	0.0180	0.0002
75	15	ENV	83	PASTNROSGR	A	R	10					
90	18	POL	616	PIDWVVOGR	I	R	9	0.0002	0.0005			
80	16	POL	496	PILGFRK	I	K	8					
95	19	POL	20	PLEEELR	L	R	B	0.0002	-0.0002	-0.0009	-0.0009	0.0001
100	20	POL	2	PLSYCHFR	L	R	B	-0.0002	-0.0002	-0.0009	-0.0009	0.0001
75	15	POL	2	PLSYCHFR	L	K	9	0.0011	0.0031	0.0006	0.0008	0.0002
85	17	POL	98	PLTVNEGR	L	R	B	0.0002	-0.0002	-0.0009	-0.0009	0.0001
85	17	POL	98	PLTVNEGR	L	R	9	0.0008	0.0005	0.0004	0.0027	0.0002
90	18	X	20	PVGAESPGR	V	R	9	0.0002	0.0005	0.0004	0.0043	0.0002
85	17	POL	612	PVNPIDWK	V	K	9	0.0310	0.1400	0.0002	0.0006	0.0009
95	19	POL	654	QAFISPTVK	A	K	10	0.0450	0.5400	0.0010	0.0057	1.2000
80	16	ENV	179	QAGFTLIR	A	R	9					
75	15	NJC	168	OSFRNRNOSR	A	R	11					
80	16	NJC	189	OSSQLSH	S	R	B					
75	15	POL	106	RLKLMPAR	L	R	9	0.0850	0.0002	3.1000	0.0490	0.0002
75	15	X	128	RLKVMALGCR	L	R	11					
95	19	POL	376	RLWDSQFSR	L	R	11	0.2800	3.8000	2.6000	1.2000	6.1000
95	19	NJC	183	RSPRRITPSR	S	R	11	-0.0007	-0.0003	0.0190	-0.0023	0.0003
75	15	NJC	167	RSCSPRRR	S	R	9					
95	18	NJC	188	RTPSPRRR	T	R	8	-0.0002	0.0005	0.0033	0.0014	0.0002
75	15	NJC	167	RSCSPRRR	S	R	9	0.0054	0.0005	0.0002	0.0016	0.0003
95	18	NJC	188	RTPSPRRR	T	R	9	0.0190	0.0290	-0.0002	-0.0003	0.0001
100	20	POL	357	RVTGVALVDK	V	K	11	-0.0002	0.0020	0.0029	0.0024	0.0360
90	18	X	65	SAGPCALR	A	R	8	-0.0002	0.0071	0.0280	0.0081	0.0690
95	19	POL	520	SAICSVRR	A	R	9	0.0058	0.2100	0.0150	0.0650	0.3800
90	16	POL	771	SALNPADPSR	A	R	11	-0.0004	-0.0003	-0.0012	-0.0023	0.0003
75	15	POL	585	SLGHLNPNK	L	R	10					
90	18	X	84	SSAGPCALR	S	R	9	0.0080	0.1400	0.3300	0.1600	0.7500
95	18	NJC	170	STUPETVRR	T	R	10	0.0007	0.0600	0.0080	0.0240	0.0250
95	18	NJC	170	STUPETVRR	T	R	11	0.0150	1.4000	0.1000	0.1600	0.3100
80	16	ENV	85	STNRKSGR	T	R	8					
75	15	X	104	STTDLEAVFK	T	K	10	0.0066	2.7000			
85	17	POL	716	TAELLAACFAR	A	R	11	0.0008	0.0023	0.0068	0.1600	0.0590
95	18	NJC	171	TLPETVRR	L	R	9	0.0008	0.0002	0.0008	0.0024	0.0180
95	18	NJC	171	TLPETVRR	L	R	10	0.0007	0.0230	0.0006	0.0120	0.0440
95	19	NJC	171	TLPETVRR	L	R	11	0.0005	0.0160	0.0061	0.0710	0.6400
100	20	POL	150	TLWKAGILYK	L	K	10	5.3000	0.3600	0.0051	0.0010	0.0130
100	20	POL	150	TLWKAGILYK	L	R	11	0.0082	0.0095	0.1000	0.1100	0.0840
95	18	POL	519	TSACSVRR	S	R	9	0.0018	0.0006	0.0600	0.0820	0.0048
95	19	POL	519	TSACSVRR	S	R	10	0.0018	0.0006	0.0030	0.0068	0.0048
75	15	X	105	TTDLEAVFK	T	K	9	0.0006	0.9200	0.0006	0.0012	0.0170
75	15	ENV	278	TTSTGPK	T	K	8					
80	16	NJC	175	TVWRPGR	T	R	9	0.0008	0.0005	0.2500	0.1400	0.0095
80	16	NJC	176	TVWRPGR	T	R	8	0.0003	0.0001			
80	16	NJC	176	TVWRPGR	T	R	11					

0.0001 0.0002 0.0003 0.0004 0.0005 0.0006 0.0007 0.0008 0.0009 0.0010 0.0011 0.0012 0.0013 0.0014 0.0015 0.0016 0.0017 0.0018 0.0019 0.0020 0.0021 0.0022 0.0023 0.0024 0.0025 0.0026 0.0027 0.0028 0.0029 0.0030 0.0031 0.0032 0.0033 0.0034 0.0035 0.0036 0.0037 0.0038 0.0039 0.0040 0.0041 0.0042 0.0043 0.0044 0.0045 0.0046 0.0047 0.0048 0.0049 0.0050 0.0051 0.0052 0.0053 0.0054 0.0055 0.0056 0.0057 0.0058 0.0059 0.0060 0.0061 0.0062 0.0063 0.0064 0.0065 0.0066 0.0067 0.0068 0.0069 0.0070 0.0071 0.0072 0.0073 0.0074 0.0075 0.0076 0.0077 0.0078 0.0079 0.0080 0.0081 0.0082 0.0083 0.0084 0.0085 0.0086 0.0087 0.0088 0.0089 0.0090 0.0091 0.0092 0.0093 0.0094 0.0095 0.0096 0.0097 0.0098 0.0099 0.0100 0.0101 0.0102 0.0103 0.0104 0.0105 0.0106 0.0107 0.0108 0.0109 0.0110 0.0111 0.0112 0.0113 0.0114 0.0115 0.0116 0.0117 0.0118 0.0119 0.0120 0.0121 0.0122 0.0123 0.0124 0.0125 0.0126 0.0127 0.0128 0.0129 0.0130 0.0131 0.0132 0.0133 0.0134 0.0135 0.0136 0.0137 0.0138 0.0139 0.0140 0.0141 0.0142 0.0143 0.0144 0.0145 0.0146 0.0147 0.0148 0.0149 0.0150 0.0151 0.0152 0.0153 0.0154 0.0155 0.0156 0.0157 0.0158 0.0159 0.0160 0.0161 0.0162 0.0163 0.0164 0.0165 0.0166 0.0167 0.0168 0.0169 0.0170 0.0171 0.0172 0.0173 0.0174 0.0175 0.0176 0.0177 0.0178 0.0179 0.0180 0.0181 0.0182 0.0183 0.0184 0.0185 0.0186 0.0187 0.0188 0.0189 0.0190 0.0191 0.0192 0.0193 0.0194 0.0195 0.0196 0.0197 0.0198 0.0199 0.0200 0.0201 0.0202 0.0203 0.0204 0.0205 0.0206 0.0207 0.0208 0.0209 0.0210 0.0211 0.0212 0.0213 0.0214 0.0215 0.0216 0.0217 0.0218 0.0219 0.0220 0.0221 0.0222 0.0223 0.0224 0.0225 0.0226 0.0227 0.0228 0.0229 0.0230 0.0231 0.0232 0.0233 0.0234 0.0235 0.0236 0.0237 0.0238 0.0239 0.0240 0.0241 0.0242 0.0243 0.0244 0.0245 0.0246 0.0247 0.0248 0.0249 0.0250 0.0251 0.0252 0.0253 0.0254 0.0255 0.0256 0.0257 0.0258 0.0259 0.0260 0.0261 0.0262 0.0263 0.0264 0.0265 0.0266 0.0267 0.0268 0.0269 0.0270 0.0271 0.0272 0.0273 0.0274 0.0275 0.0276 0.0277 0.0278 0.0279 0.0280 0.0281 0.0282 0.0283 0.0284 0.0285 0.0286 0.0287 0.0288 0.0289 0.0290 0.0291 0.0292 0.0293 0.0294 0.0295 0.0296 0.0297 0.0298 0.0299 0.0300 0.0301 0.0302 0.0303 0.0304 0.0305 0.0306 0.0307 0.0308 0.0309 0.0310 0.0311 0.0312 0.0313 0.0314 0.0315 0.0316 0.0317 0.0318 0.0319 0.0320 0.0321 0.0322 0.0323 0.0324 0.0325 0.0326 0.0327 0.0328 0.0329 0.0330 0.0331 0.0332 0.0333 0.0334 0.0335 0.0336 0.0337 0.0338 0.0339 0.0340 0.0341 0.0342 0.0343 0.0344 0.0345 0.0346 0.0347 0.0348 0.0349 0.0350 0.0351 0.0352 0.0353 0.0354 0.0355 0.0356 0.0357 0.0358 0.0359 0.0360 0.0361 0.0362 0.0363 0.0364 0.0365 0.0366 0.0367 0.0368 0.0369 0.0370 0.0371 0.0372 0.0373 0.0374 0.0375 0.0376 0.0377 0.0378 0.0379 0.0380 0.0381 0.0382 0.0383 0.0384 0.0385 0.0386 0.0387 0.0388 0.0389 0.0390 0.0391 0.0392 0.0393 0.0394 0.0395 0.0396 0.0397 0.0398 0.0399 0.0400 0.0401 0.0402 0.0403 0.0404 0.0405 0.0406 0.0407 0.0408 0.0409 0.0410 0.0411 0.0412 0.0413 0.0414 0.0415 0.0416 0.0417 0.0418 0.0419 0.0420 0.0421 0.0422 0.0423 0.0424 0.0425 0.0426 0.0427 0.0428 0.0429 0.0430 0.0431 0.0432 0.0433 0.0434 0.0435 0.0436 0.0437 0.0438 0.0439 0.0440 0.0441 0.0442 0.0443 0.0444 0.0445 0.0446 0.0447 0.0448 0.0449 0.0450 0.0451 0.0452 0.0453 0.0454 0.0455 0.0456 0.0457 0.0458 0.0459 0.0460 0.0461 0.0462 0.0463 0.0464 0.0465 0.0466 0.0467 0.0468 0.0469 0.0470 0.0471 0.0472 0.0473 0.0474 0.0475 0.0476 0.0477 0.0478 0.0479 0.0480 0.0481 0.0482 0.0483 0.0484 0.0485 0.0486 0.0487 0.0488 0.0489 0.0490 0.0491 0.0492 0.0493 0.0494 0.0495 0.0496 0.0497 0.0498 0.0499 0.0500 0.0501 0.0502 0.0503 0.0504 0.0505 0.0506 0.0507 0.0508 0.0509 0.0510 0.0511 0.0512 0.0513 0.0514 0.0515 0.0516 0.0517 0.0518 0.0519 0.0520 0.0521 0.0522 0.0523 0.0524 0.0525 0.0526 0.0527 0.0528 0.0529 0.0530 0.0531 0.0532 0.0533 0.0534 0.0535 0.0536 0.0537 0.0538 0.0539 0.0540 0.0541 0.0542 0.0543 0.0544 0.0545 0.0546 0.0547 0.0548 0.0549 0.0550 0.0551 0.0552 0.0553 0.0554 0.0555 0.0556 0.0557 0.0558 0.0559 0.0560 0.0561 0.0562 0.0563 0.0564 0.0565 0.0566 0.0567 0.0568 0.0569 0.0570 0.0571 0.0572 0.0573 0.0574 0.0575 0.0576 0.0577 0.0578 0.0579 0.0580 0.0581 0.0582 0.0583 0.0584 0.0585 0.0586 0.0587 0.0588 0.0589 0.0590 0.0591 0.0592 0.0593 0.0594 0.0595 0.0596 0.0597 0.0598 0.0599 0.0600 0.0601 0.0602 0.0603 0.0604 0.0605 0.0606 0.0607 0.0608 0.0609 0.0610 0.0611 0.0612 0.0613 0.0614 0.0615 0.0616 0.0617 0.0618 0.0619 0.0620 0.0621 0.0622 0.0623 0.0624 0.0625 0.0626 0.0627 0.0628 0.0629 0.0630 0.0631 0.0632 0.0633 0.0634 0.0635 0.0636 0.0637 0.0638 0.0639 0.0640 0.0641 0.0642 0.0643 0.0644 0.0645 0.0646 0.0647 0.0648 0.0649 0.0650 0.0651 0.0652 0.0653 0.0654 0.0655 0.0656 0.0657 0.0658 0.0659 0.0660 0.0661 0.0662 0.0663 0.0664 0.0665 0.0666 0.0667 0.0668 0.0669 0.0670 0.0671 0.0672 0.0673 0.0674 0.0675 0.0676 0.0677 0.0678 0.0679 0.0680 0.0681 0.0682 0.0683 0.0684 0.0685 0.0686 0.0687 0.0688 0.0689 0.0690 0.0691 0.0692 0.0693 0.0694 0.0695 0.0696 0.0697 0.0698 0.0699 0.0700 0.0701 0.0702 0.0703 0.0704 0.0705 0.0706 0.0707 0.0708 0.0709 0.0710 0.0711 0.0712 0.0713 0.0714 0.0715 0.0716 0.0717 0.0718 0.0719 0.0720 0.0721 0.0722 0.0723 0.0724 0.0725 0.0726 0.0727 0.0728 0.0729 0.0730 0.0731 0.0732 0.0733 0.0734 0.0735 0.0736 0.0737 0.0738 0.0739 0.0740 0.0741 0.0742 0.0743 0.0744 0.0745 0.0746 0.0747 0.0748 0.0749 0.0750 0.0751 0.0752 0.0753 0.0754 0.0755 0.0756 0.0757 0.0758 0.0759 0.0760 0.0761 0.0762 0.0763 0.0764 0.0765 0.0766 0.0767 0.0768 0.0769 0.0770 0.0771 0.0772 0.0773 0.0774 0.0775 0.0776 0.0777 0.0778 0.0779 0.0780 0.0781 0.0782 0.0783 0.0784 0.0785 0.0786 0.0787 0.0788 0.0789 0.0790 0.0791 0.0792 0.0793 0.0794 0.0795 0.0796 0.0797 0.0798 0.0799 0.0800 0.0801 0.0802 0.0803 0.0804 0.0805 0.0806 0.0807 0.0808 0.0809 0.0810 0.0811 0.0812 0.0813 0.0814 0.0815 0.0816 0.0817 0.0818 0.0819 0.0820 0.0821 0.0822 0.0823 0.0824 0.0825 0.0826 0.0827 0.0828 0.0829 0.0830 0.0831 0.0832 0.0833 0.0834 0.0835 0.0836 0.0837 0.0838 0.0839 0.0840 0.0841 0.0842 0.0843 0.0844 0.0845 0.0846 0.0847 0.0848 0.0849 0.0850 0.0851 0.0852 0.0853 0.0854 0.0855 0.0856 0.0857 0.0858 0.0859 0.0860 0.0861 0.0862 0.0863 0.0864 0.0865 0.0866 0.0867 0.0868 0.0869 0.0870 0.0871 0.0872 0.0873 0.0874 0.0875 0.0876 0.0877 0.0878 0.0879 0.0880 0.0881 0.0882 0.0883 0.0884 0.0885 0.0886 0.0887 0.0888 0.0889 0.0890 0.0891 0.0892 0.0893 0.0894 0.0895 0.0896 0.0897 0.0898 0.0899 0.0900 0.0901 0.0902 0.0903 0.0904 0.0905 0.0906 0.0907 0.0908 0.0909 0.0910 0.0911 0.0912 0.0913 0.0914 0.0915 0.0916 0.0917 0.0918 0.0919 0.0920 0.0921 0.0922 0.0923 0.0924 0.0925 0.0926 0.0927 0.0928 0.0929 0.0930 0.0931 0.0932 0.0933 0.0934 0.0935 0.0936 0.0937 0.0938 0.0939 0.0940 0.0941 0.0942 0.0943 0.0944 0.0945 0.0946 0.0947 0.0948 0.0949 0.0950 0.0951 0.0952 0.0953 0.0954 0.0955 0.0956 0.0957 0.0958 0.0959 0.0960 0.0961 0.0962 0.0963 0.0964 0.0965 0.0966 0.0967 0.0968 0.0969 0.0970 0.0971 0.0972 0.0973 0.0974 0.0975 0.0976 0.0977 0.0978 0.0979 0.0980 0.0981 0.0982

# HBV A03 SUPER MOTIF (with binding information)

Conservancy	Frequency	Protein	Position	Sequence	P2	C-term	AA	A*0301	A*1101	A*3101	A*3301	A*6801
90	18	X	133	WGGGRHK	L	K	8	0.0150	0.0002	-0.0005	-0.0009	0.0001
80	16	EW	177	WLAGFLLTR	L	R	11	0.0040	0.0290	0.0750	0.0270	0.0360
80	18	NJC	120	VDFGWMR	S	R	8	0.0130	0.0170	0.0031	0.0013	0.0004
90	20	POL	48	VSIPVTHK	S	K	8	0.0390	0.0920	0.0002	0.0006	0.0022
100	20	POL	358	VTGGVFLVDK	T	K	10	0.0015	0.0750	0.0013	0.0170	0.0330
100	20	POL	378	WDFSGFSR	V	R	9	0.0027	0.0001			
100	20	POL	177	WRRRGSRPR	V	R	10					
80	16	NJC	177	WRRRGSRPR	V	R	11	0.0008	0.0005	0.0020	0.0052	0.0001
80	16	NJC	177	WRRRGSRPR	V	R	9	-0.0002	0.0005	-0.0003	0.0039	0.0490
95	19	NJC	125	WRTTPAYR	L	K	8	0.0030	0.0013	-0.0006	0.0006	0.0002
90	18	POL	314	WLFNNK	L	K	11	0.0001	0.0001			
85	17	NJC	26	WLVGMDPYK	L	K	9	0.0005	0.0002			
100	20	POL	122	YLPDGGIK	L	R	10	0.0005	0.0043	0.0002	0.0006	0.0001
90	18	NJC	118	YLVDFGWMR	L	R	10					
90	18	POL	538	YMDVVLGAK	M	K	10					
80	16	POL	493	YSPILGFR	S	R	10					
80	16	POL	493	YSPILGFR	S	K	11					

HBV A03 SUPER MOTIF (with binding information)

Table X

## HBV A24 SUPER MOTIF (With binding information)

A\*2401

Conservancy	Freq	Protein	Position	Sequence	String	
					XFXXXXXF	
95	19	POL	529	AFPHCLAF	XFXXXXXXFY	
95	19	POL	529	AFPHCLAFSY	XFXXXXXXXM	0.0012
95	19	POL	529	AFPHCLAFSYM	XFXXXXXXXL	0.0009
95	19	X	82	AFSSAGPCAL	XFXXXXXXXL	
90	18	POL	535	AFSYMDDVVL	XFXXXXXXY	
95	19	POL	655	AFTFSPTY	XFXXXXXXXF	
95	19	POL	655	AFTFSPTYKAF	XIXXXXXXXF	
95	19	POL	521	AICSVVRRAF	XIXXXXXXL	
90	18	NUC	58	AILCWGEL	XIXXXXXXM	
90	18	NUC	58	AILCWGELM	XLXXXXXXI	
95	19	POL	642	ALMPYACI	XLXXXXXXW	
95	19	NUC	54	ALRQAILCW	XMXXXXXXF	
80	16	BW	108	AMQWNSTTF	XTXXXXXXL	
95	19	POL	690	ATPTGWGL	XTXXXXXXI	
75	15	POL	690	ATPTGWGLAI	XVXXXXXL	
95	19	POL	397	AVPNLQSL	XVXXXXXXXL	0.0260
95	19	POL	397	AVPNLQSLTNL	XYXXXXXXI	0.0220
100	20	NUC	131	AYRPPNAPI	XYXXXXXXXL	
100	20	NUC	131	AYRPPNAPIL	XFXXXXXXXI	
75	15	POL	607	CFRKLPNRPI	XXXXXXXW	
100	20	BW	312	CIPISSW	XXXXXXXXF	
100	20	BW	312	CIPISSWAF	XLXXXXXM	
85	17	NUC	23	CLGWLWGM	XLXXXXXXI	
85	17	NUC	23	CLGWLWGMCI	XLXXXXXL	
100	20	BW	253	CLIFLLVL	XLXXXXXXL	
100	20	BW	253	CLIFLLVLL	XLXXXXXXXI	
100	20	BW	253	CLIFLLVLLDY	XLXXXXXXXI	
95	19	BW	239	CLRRFIIF	XLXXXXXXF	
95	19	BW	239	CLRRFIIFL	XLXXXXXXF	
95	19	BW	239	CLRRFIIFL	XLXXXXXXXI	
75	15	BW	239	CLRRFIIFL	XTXXXXXXXW	
75	15	BW	239	CTCIPISW	XIXXXXXF	
100	20	NUC	31	DIDPYKEF	XLXXXXXXL	
90	18	NUC	29	DLLDTASAL	XLXXXXXXXI	
85	17	NUC	29	DLLDTASALY	XLXXXXXXXI	
85	17	NUC	29	DLLGGLNLSI	XLXXXXXXXI	
95	19	POL	40	DTASALYREAL	XTXXXXXXXL	
80	16	NUC	32	DMKVCQRI	XVXXXXXI	
85	17	POL	618	DMKVCQRI	XVXXXXXXXL	0.0002
85	17	POL	618	DMKVCQRI	XYXXXXXXXL	
90	18	BW	262	DYQGMPLVCL	XLXXXXXL	
80	16	X	122	ELGEEIL	XLXXXXXXF	
95	19	NUC	43	ELLSFLPSOF	XLXXXXXXF	
95	19	NUC	43	ELLSFLPSOFF	XYXXXXXXW	0.0340
90	18	NUC	117	EYLVSGWV	XYXXXXXXI	
90	18	NUC	117	EYLVSGWVM	XFXXXXXI	
100	20	BW	382	FFCLWYI	XFXXXXXL	
80	16	BW	182	FFLLTRIL	XFXXXXXXI	
80	16	BW	182	FFLLTRILTI	XFXXXXXXXF	
85	17	BW	13	FFPOHOLPAF	XIXXXXXL	
80	16	BW	243	FIIFLFI	XIXXXXXL	
80	16	BW	243	FIIFLFI	XIXXXXXL	
80	16	BW	243	FIIFLFI	XIXXXXXL	
80	16	BW	248	FILLLCU	XIXXXXXI	
80	16	BW	248	FILLLCIF	XIXXXXXF	
80	16	BW	248	FILLLCIFL	XIXXXXXL	
80	16	BW	248	FILLLCIFL	XIXXXXXL	
80	16	BW	248	FILLLCIFL	XLXXXXXL	
80	16	BW	246	FLRILLCL	XLXXXXXXI	
80	16	BW	246	FLRILLCL	XLXXXXXXF	
80	16	BW	246	FLRILLCL	XLXXXXXL	
75	15	BW	171	FLGPLLV	XLXXXXXXI	
95	19	POL	513	FLLAQFTSAI	XLXXXXXXI	
95	19	POL	562	FLSLGIHL	XLXXXXXXI	
95	19	BW	183	FLTRILTI	XLXXXXXXI	
80	16	BW	258	FLVLLDY	XLXXXXXXY	
95	19	BW	256	FLVLLDYQGM	XLXXXXXXXM	
95	19	POL	656	FTFSPTYKAF	XTXXXXXXXF	
95	19	POL	656	FTFSPTYKAF	XTXXXXXXXL	
95	19	POL	635	FTQCGYPAL	XTXXXXXXXI	
95	19	POL	635	FTQCGYPALM	XTXXXXXXXM	
95	19	BW	346	FVGLSPTVW	XVXXXXXXW	
95	19	BW	346	FVGLSPTVW	XVXXXXXXXL	
90	18	X	132	FVGGCPHKL	XVXXXXXXL	
95	19	BW	342	FVQWVGL	XVXXXXXL	
90	18	POL	766	FVYVSAL	XVXXXXXL	
95	19	POL	630	GFAAPFTQCGY	XFXXXXXXI	
80	16	BW	181	GFLLTRI	XFXXXXXXL	
80	16	BW	181	GFLLTRIL	XFXXXXXXI	
80	16	BW	181	GFLLTRILTI	XFXXXXXXXI	

## HBV A24 SUPER MOTIF (With binding information)

A\*2401

Conservancy	Freq	Protein	Position	Sequence	String	
95	19	BN	12	GFFPDHQL	XFXXXXXL	
75	15	BN	170	GFLGPLVL	XFXXXXXL	
80	16	POL	500	GFRKIPMGVGL	XFXXXXXXXL	
95	19	POL	627	GLLGFAAPF	XLXXXXXXF	
95	19	POL	509	GLSPFLACF	XLXXXXXXF	
100	20	BN	348	GLSPTVWL	XLXXXXXL	
75	15	BN	348	GLSPTVWLSVI	XLXXXXXXXI	
85	17	NUC	29	GMDIDPYKEF	XMXXXXXXF	
90	18	BN	285	GMLPVCPL	XMXXXXXL	
90	18	POL	735	GTDSNVL	XTXXXXXXL	
75	15	BN	13	GTNLSVFNPL	XTXXXXXXXL	
80	16	POL	783	GTSFVYVPSAL	XTXXXXXXXL	
80	16	POL	507	GVGLSPFL	XVXXXXXL	
90	16	POL	507	GVGLSPFL	XVXXXXXXL	
85	17	NUC	123	GWIRTTPAY	XVXXXXXXY	
85	17	NUC	25	GWLWGMCI	XWXXXXXI	
85	17	NUC	25	GWLWGMIDPY	XWXXXXXXXY	0.0024
85	17	BN	65	GWSPQACGI	XWXXXXXI	0.0003
85	17	BN	65	GWSPQACGIL	XWXXXXXXL	
95	19	POL	639	GYPALMPL	XYXXXXXL	0.0490
95	19	POL	639	GYPALMPY	XYXXXXXXY	0.0110
95	19	BN	234	GYRWMCLRRF	XYXXXXXXF	
95	19	BN	234	GYRWMCLRRFI	XYXXXXXXXI	
85	17	POL	579	GYSLNRMGY	XYXXXXXXY	0.0002
75	15	POL	579	GYSLNRMGYVI	XYXXXXXXXI	
80	16	POL	820	HFASPLHVAW	XFXXXXXXW	
75	15	POL	7	HFRKLLL	XFXXXXXL	
80	16	POL	435	HLLVGSSGL	XLXXXXXXL	
75	15	POL	589	HLNPNKTKRW	XLXXXXXXW	
80	16	POL	491	HLYSHPII	XLXXXXXI	
80	16	POL	491	HLYSHPIIL	XLXXXXXXL	
80	16	POL	491	HLYSHPIILGF	XLXXXXXXXF	
85	17	POL	715	HTAELLAACF	XTXXXXXXF	
100	20	NUC	52	HTALRQAI	XTXXXXXI	
95	19	NUC	52	HTALRQAIL	XTXXXXXXL	
95	19	NUC	52	HTALRQAILCW	XTXXXXXXXW	
100	20	POL	149	HTLWKAGI	XTXXXXXI	
100	20	POL	149	HTLWKAGIL	XTXXXXXXL	
100	20	POL	149	HTLWKAGILY	XTXXXXXXY	
100	20	POL	146	HYLTLWKAGI	XYXXXXXXXI	
100	20	BN	381	IFFCLWVY	XFXXXXXXY	0.0087
100	20	BN	381	IFFCLWVYI	XFXXXXXXI	
80	16	BN	245	IFLFILL	XFXXXXXL	
80	16	BN	245	IFLFILLCL	XFXXXXXXL	
80	16	BN	245	IFLFILLCLJ	XFXXXXXXXI	
95	19	BN	255	IFLLVLDY	XFXXXXXXY	
80	18	BN	244	IIFLFIIL	XIXXXXXL	
80	16	BN	244	IIFLFIIL	XIXXXXXL	
80	16	BN	244	IIFLFIILCL	XIXXXXXXXL	
80	16	POL	497	IILGFRKI	XIXXXXXI	
80	16	POL	497	IILGFRKIPM	XIXXXXXXXM	
80	16	NUC	59	ILCWGELM	XLXXXXXM	
80	16	POL	498	ILGFRKIPM	XLXXXXXXM	
100	20	BN	249	ILLCLIF	XLXXXXXF	
100	20	BN	249	ILLCLIFL	XLXXXXXXL	
100	20	BN	249	ILLCLIFLL	XLXXXXXXXL	
80	16	POL	760	ILRGTSFVY	XLXXXXXXY	
95	19	BN	188	ILTIQSL	XLXXXXXL	
90	18	BN	188	ILTIQSLDSW	XLXXXXXXXW	
90	18	POL	625	IVGLLGFAAPF	XVXXXXXXF	0.0004
85	17	BN	358	IWMWVYWGFS	XWXXXXXXXL	0.0020
95	19	POL	395	KFAVPLQSL	XFXXXXXXL	
80	16	POL	503	KIPMGVGL	XIXXXXXL	
80	16	POL	503	KIPMGVGLSPF	XIXXXXXXXF	
85	17	NUC	21	KLCGLWLW	XLXXXXXXW	
85	17	NUC	21	KLCGLWLWGM	XLXXXXXXM	
95	19	POL	489	KLHLYSHPI	XLXXXXXXI	
80	16	POL	489	KLHLYSHPII	XLXXXXXXI	
80	16	POL	489	KLHLYSHPIIL	XLXXXXXXXL	
75	15	POL	108	KLIMPARF	XLXXXXXF	
75	15	POL	108	KLIMPARFY	XLXXXXXXY	
80	16	POL	610	KLPVNRPI	XLXXXXXI	
80	16	POL	610	KLPVNRPIDW	XLXXXXXXXW	
95	19	POL	574	KTKRWGYSL	XTXXXXXXL	
85	17	POL	574	KTKRWGYSUNF	XTXXXXXXXF	
85	17	POL	620	KVCQRIVGL	XVXXXXXXL	
85	17	POL	620	KVCQRIVGLL	XVXXXXXXXL	
95	19	POL	55	KVGNFTGL	XVXXXXXL	

## HBV A24 SUPER MOTIF (With binding information)

A\*2401

Conservancy	Freq	Protein	Position	Sequence	String	Pptide	
					XVXXXXXX	1.0166	
95	19	POL	55	KVGNFTGLY	XVXXXXXX		
85	17	X	91	KVLHKRTL	XVXXXXXX	1.0800	0.0028
85	17	X	91	KVLHKRTLGL	XVXXXXXX	5.0063	
100	20	POL	121	KYLPDKGI	XYXXXXXX	17.0132	3.6000
85	17	POL	745	KYTSFPWL	XYXXXXXX	2.0061	
85	17	POL	745	KYTSFPWLL	XYXXXXXX	17.0247	
80	16	BN	247	LFILLCL	XFXXXXXX		
80	16	BN	247	LFILLCLI	XFXXXXXX		
80	16	BN	247	LFILLCLIF	XFXXXXXX		
80	16	BN	247	LFILLCLIFL	XFXXXXXX		
80	16	BN	254	LIPLVL	XIXXXXXX	Chisari 4.014	
100	20	BN	254	LIPLVL	XIXXXXXXX	1.0899	
95	19	POL	109	LIMPARFY	XIXXXXXXX	26.0028	
100	20	POL	514	LLAQFTSAI	XLXXXXXX	3.0010	
95	19	BN	251	LLCUIFL	XLXXXXXX	Chisari 4.015	
100	20	BN	251	LLCUIFLVL	XLXXXXXX	1.0898	
100	20	BN	251	LLCUIFLVLL	XLXXXXXX	Chisari 4.016	
100	20	NUC	30	LLDTASAL	XLXXXXXX	1.0155	
85	17	NUC	30	LLDTASALY	XLXXXXXX	Chisari 4.021	
85	17	BN	260	LLDYOGML	XLXXXXXX		
95	19	POL	752	LLGCAANW	XLXXXXXX	3.0013	
80	16	POL	752	LLGCAANWI	XLXXXXXX	1.0912	
80	16	POL	752	LLGCAANWIL	XLXXXXXX		
95	19	POL	628	LLGFAAPF	XLXXXXXX		
75	15	BN	63	LLGWSPOAGI	XLXXXXXX	Chisari 4.017	
100	20	BN	250	LLCUIFL	XLXXXXXX	1.0834	
100	20	BN	250	LLCUIFL	XLXXXXXX	Chisari 4.018	
100	20	BN	250	LLCUIFLVL	XLXXXXXX	17.0112	
100	20	BN	378	LLPIFFCL	XLXXXXXX		
100	20	BN	378	LLPIFFCLW	XLXXXXXX	26.0549	
100	20	BN	378	LLPIFFCLWVY	XLXXXXXX		
95	19	NUC	44	LLSFLPSDF	XLXXXXXX		
95	19	NUC	44	LLSFLPSDF	XLXXXXXX		
95	19	POL	563	LLSLGIHL	XLXXXXXX		
90	18	POL	407	LLSSNLW	XLXXXXXX	1.0184	
90	18	POL	407	LLSSNLW	XLXXXXXX		
90	18	POL	407	LLSSNLWLSL	XLXXXXXX	Chisari 4.053	
80	16	BN	184	LLTRILTI	XLXXXXXX		
80	16	POL	436	LLVGSSGL	XLXXXXXX	3.0207	
80	16	BN	257	LLVLDYOGM	XLXXXXXX		
95	19	BN	257	LLVLDYOGML	XLXXXXXX		
95	19	BN	175	LLVLQAGF	XLXXXXXX	20.0121	
95	19	BN	175	LLVLQAGFF	XLXXXXXX	1.0892	
90	18	BN	175	LLVLQAGFFL	XLXXXXXX	Chisari 4.028	
90	18	BN	175	LLVLQAGFFLL	XLXXXXXX		
100	20	BN	338	LLVPFQWF	XLXXXXXX		
100	20	BN	338	LLVPFQWF	XLXXXXXX	1.0844	
90	18	NUC	100	LLWFHISCL	XLXXXXXX		
85	17	NUC	100	LLWFHISCLTF	XLXXXXXX	17.0130	
95	19	POL	643	LMPLYACI	XTXXXXXX		
75	15	NUC	137	LTFGRETVL	XTXXXXXX		
75	15	NUC	137	LTFGRETVLEY	XTXXXXXX		
90	18	BN	189	LTIPOSLDSW	XTXXXXXX		
90	18	BN	189	LTIPOSLDSWW	XTXXXXXX		
90	18	POL	404	LTIPOSLDSW	XTXXXXXX		
90	18	POL	404	LTIPOSLDSW	XTXXXXXX		
80	16	BN	185	LTIPOSLDSW	XTXXXXXX		
85	17	POL	99	LTVNEKRL	XTXXXXXX	3.0034	
95	19	BN	258	LVLDYQGM	XVXXXXXX	1.0515	
95	19	BN	258	LVLDYQGM	XVXXXXXX		
95	19	BN	176	LVLDYQGM	XVXXXXXX	1.0827	
90	18	BN	176	LVLDYQGM	XVXXXXXX	1.0893	
90	18	BN	176	LVLDYQGM	XVXXXXXX		
100	20	BN	339	LVPPVQWF	XVXXXXXX		
95	19	BN	339	LVPPVQWF	XVXXXXXX	Chisari 4.078	
90	18	NUC	119	LVSFGWM	XVXXXXXX		
100	20	POL	377	LWDFSQF	XVXXXXXX		
90	18	NUC	101	LWFHISCL	XVXXXXXX	26.0373	
85	17	NUC	101	LWFHISCLTF	XVXXXXXX		
85	17	NUC	27	LWGMIDPY	XVXXXXXX		
100	20	POL	151	LWKAGILY	XVXXXXXX		1.1000
80	16	POL	492	LYSHPIIL	XVXXXXXX	2.0181	0.0012
80	16	POL	492	LYSHPIILGF	XVXXXXXX	1.0839	0.0001
85	17	BN	360	MMWYWGPSL	XVXXXXXX	1039.01	
85	17	BN	360	MMWYWGPSLY	XVXXXXXX	17.0249	
85	17	BN	361	MMWYWGPSL	XVXXXXXX	1039.02	0.0027
85	17	BN	361	MMWYWGPSLY	XVXXXXXX		
95	19	POL	561	NFLSLGI	XFXXXXXX		

## HBV A24 SUPER MOTIF (With blinding information)

A\*2401

Conservancy	Freq	Protein	Position	Sequence	String
					0.0099
95	19	POL	561	NFLSLGIHL	XFXXXXXXXXL
95	19	POL	42	NLGNLNVSI	XLXXXXXXXXI
95	19	POL	42	NLGNLNVSIW	XLXXXXXXXXW
90	18	POL	406	NLLSSNLSW	XLXXXXXXXXW
90	18	POL	406	NLLSSNLSWL	XLXXXXXXXXL
95	19	POL	45	NLNVSIW	XLXXXXXW
100	20	POL	400	NLOSLTNL	XLXXXXXL
100	20	POL	400	NLOSLTNLL	XLXXXXXL
75	15	BNV	15	NLSVNPPL	XLXXXXXXF
75	15	BNV	15	NLSVNPPLGF	XWXXXXXXF
80	16	POL	758	NWLRTGSF	XWXXXXXXY
80	16	POL	758	NWLRTGSFVY	XFXXXXXXXXI
95	19	POL	512	PFLAQFTSAI	XFXXXXXXXXL
95	19	POL	534	PFTQCGYPAL	XFXXXXXXXXM
95	19	POL	534	PFTQCGYPALM	XFXXXXXXXXM
95	19	BNV	341	PFQWVQL	XFXXXXXXXXI
85	17	BNV	380	PIFWKVCRI	XIXXXXXXX
100	20	BNV	380	PIFFCLWVY	XIXXXXXXXI
100	20	BNV	380	PIFFCLWVYI	XIXXXXXXL
85	17	POL	713	PIHTAELL	XIXXXXXXI
80	16	POL	498	PIILGFRKI	XIXXXXXXXM
80	16	POL	498	PIILGFRKIPM	XIXXXXXX
100	20	BNV	314	PISSWAF	XLXXXXXXY
100	20	POL	124	PLDKGKPY	XLXXXXXXY
100	20	POL	124	PLDKGKPY	XLXXXXXL
95	19	BNV	20	PLLELPRL	XLXXXXXXL
95	19	BNV	10	PLGFFPHQL	XLXXXXXXL
100	20	POL	427	PLHPAAMPIL	XLXXXXXXL
100	20	POL	427	PLHPAAMPILL	XLXXXXXXL
100	20	BNV	377	PLPIFFCL	XLXXXXXXW
100	20	BNV	377	PLPIFFCLW	XLXXXXXXF
95	19	BNV	174	PLVLQAGF	XLXXXXXXF
95	19	BNV	174	PLVLQAGFF	XLXXXXXXL
90	18	BNV	174	PLVLQAGFLL	XLXXXXXXL
80	16	POL	711	PLPIHTAEL	XLXXXXXXL
80	16	POL	711	PLPIHTAELL	XLXXXXXXL
75	15	POL	2	PLSYQHFRKL	XLXXXXXXL
75	15	POL	2	PLSYQHFRKLL	XLXXXXXXL
85	17	POL	98	PLTVNEKRL	XWXXXXXXF
80	16	POL	505	PMGVGLSPF	XWXXXXXXL
80	16	POL	505	PMGVGLSPFL	XWXXXXXXL
80	16	POL	505	PMGVGLSPFL	XWXXXXXXL
75	15	POL	692	PTGWGLAI	XTXXXXXI
85	17	POL	797	PTTGRTSL	XTXXXXXL
85	17	POL	797	PTTGRTSLY	XTXXXXXXY
80	16	NLC	15	PTVOASKL	XTXXXXXXL
80	16	NLC	15	PTVOASKLCL	XTXXXXXI
75	15	BNV	351	PTVWLSVI	XTXXXXXXW
75	15	BNV	351	PTVWLSVIW	XTXXXXXXM
75	15	BNV	351	PTVWLSVIWM	XVXXXXXW
85	17	POL	612	PVNRPIDW	XWXXXXXXW
80	16	POL	750	PWLLGCAANW	XWXXXXXXI
80	16	POL	750	PWLLGCAANWI	XWXXXXXXI
100	20	POL	51	PWTHKVGNF	XWXXXXXXL
80	16	X	8	QLDPARDVL	XLXXXXXXL
80	16	X	8	QLDPARDVLC	XLXXXXXXL
90	18	NLC	99	QLLWPHISL	XVXXXXXXW
95	19	POL	585	QVFADATPTGW	XWXXXXXXF
95	19	BNV	344	QWVGLSPTWV	XFXXXXXI
75	15	BNV	242	RFIIFLI	XFXXXXXL
75	15	BNV	242	RFIIFLI	XFXXXXXXL
75	15	BNV	242	RFIIFLI	XFXXXXXXL
75	15	BNV	242	RFIIFLI	XFXXXXXXL
100	20	BNV	332	RFSWLSL	XFXXXXXXF
100	20	BNV	332	RFSWLSLVPF	XIXXXXXXL
80	16	BNV	167	RULTIPQSL	XDXXXXX
80	16	POL	624	RVLGLGF	XLXXXXXXF
75	15	POL	106	RLKLIMPARF	XLXXXXXXY
75	15	POL	106	RLKLIMPARFY	XLXXXXXXF
95	19	POL	376	RLVQFSQF	XTXXXXXXF
90	18	POL	353	RTPARVTGGYF	XVXXXXXL
95	19	POL	36	RVAEDLNL	XVXXXXXXL
90	18	POL	36	RVAEDNLGNL	XVXXXXXL
80	16	POL	618	RVHFASPL	XVXXXXXL
100	20	POL	357	RVTGGVFL	XWXXXXXXF
85	17	POL	577	RWGYSLNF	XWXXXXXXM
85	17	POL	577	RWGYSLNFM	XWXXXXXXY
85	17	POL	577	RWGYSLNFMGY	XWXXXXXXY

0.0002

0.0003

0.0290

## HBV A24 SUPER MOTIF (With binding information)

Conservancy	Freq	Protein	Position	Sequence	String	A*2401
95	19	EW	238	RWMCLRRF	XWXXXXXF	
95	19	EW	236	RWMCLRRF	XWXXXXXXL	0.0710
95	19	EW	236	RWMCLRRFI	XWXXXXXXI	1.1000
95	19	EW	236	RWMCLRRFIF	XWXXXXXXXF	
100	20	POL	167	SFCGSPYSW	XFXXXXXXW	0.0710
95	19	NUC	46	SFLPSOFF	XFXXXXXF	
80	18	POL	755	SFVYVPSAL	XFXXXXXXL	
100	20	POL	49	SIPWTHKVGNF	XLXXXXXXXF	
95	19	EW	194	SLDSWWTSL	XLXXXXXXL	
95	19	EW	194	SLDSWWTSLNF	XLXXXXXXXF	
95	19	POL	416	SLDVSAAF	XLXXXXXF	
95	19	POL	416	SLDVSAAFY	XLXXXXXXY	
100	20	EW	337	SLLVPFQW	XLXXXXXXW	
100	20	EW	337	SLLVPFQWF	XLXXXXXXXF	
75	15	POL	581	SLNFMGYVI	XLXXXXXXI	
95	19	X	54	SLRGLPVCAF	XLXXXXXXXF	
90	18	POL	403	SLTNLSSNL	XLXXXXXXL	
75	15	X	104	STTDLEAY	XTXXXXXXY	
75	15	X	104	STTDLEAYF	XTXXXXXXXF	
75	15	EW	17	SVPNPLGF	XVXXXXXF	
85	17	POL	548	SVCHLES	XVXXXXXL	
80	16	EW	330	SVRFWSLSL	XVXXXXXXL	
80	16	EW	330	SVRFWSLSLL	XVXXXXXXXL	
90	18	POL	739	SVVLSRKY	XVXXXXXXY	
85	17	POL	739	SVVLSRKYTSF	XVXXXXXXXF	
95	19	POL	524	SVVRRAPFHCL	XVXXXXXXXL	
95	19	POL	413	SVWLSLDVSAAF	XWXXXXXXXF	
100	20	EW	334	SWLSLLVPF	XWXXXXXXF	0.3900
95	19	POL	392	SWPKFAVPHL	XWXXXXXXXL	5.6000
100	20	EW	197	SWWTSLSNF	XWXXXXXF	
95	19	EW	197	SWWTSLSNFI	XWXXXXXXL	0.3800
90	18	POL	537	SYMDDVVL	XYXXXXXXL	
75	15	POL	4	SYCHFRKL	XYXXXXXXL	
75	15	POL	4	SYCHFRKLL	XYXXXXXXL	0.0051
75	15	POL	4	SYCHFRKLLL	XYXXXXXXL	0.0650
75	15	NUC	138	TFGRETVL	XFXXXXXL	
75	15	NUC	138	TFGRETVLEY	XFXXXXXXY	
75	15	NUC	138	TFGRETVLEYL	XFXXXXXXXL	
95	19	POL	857	TFSPITYKAF	XFXXXXXXF	0.0060
95	19	POL	857	TFSPITYKAFI	XFXXXXXXI	0.0043
90	18	EW	190	TIQSLDSW	XIXXXXXXW	
90	18	EW	190	TIQSLDSWW	XIXXXXXXXW	
100	20	POL	150	TLWKAGIL	XLXXXXXL	
100	20	POL	150	TLWKAGILY	XLXXXXXXY	
75	15	X	105	TTDLEAYF	XTXXXXXF	
85	17	POL	798	TTGRTSLY	XTXXXXXXY	
85	17	POL	100	TVNEKFR	XVXXXXXL	
80	16	NUC	16	TVQASKLCL	XVXXXXXL	
80	16	NUC	16	TVQASKLCLGW	XVXXXXXXXW	
75	15	EW	352	TVWLSVIW	XVXXXXXW	
75	15	EW	352	TVWLSVIWM	XVXXXXXXM	
95	19	POL	686	VFADATPTGW	XFXXXXXXXW	0.0180
75	15	X	131	VFLGGCRHKL	XFXXXXXXXL	
85	17	POL	543	VLGAKSVQHL	XLXXXXXXL	
90	18	X	133	VLGGCRHKL	XLXXXXXXL	
85	17	X	92	VLHKRTLGL	XLXXXXXXL	
95	19	EW	259	VLDYOGM	XLXXXXXM	
95	19	EW	259	VLDYOGML	XLXXXXXXL	
95	19	EW	177	VLDAGFRL	XLXXXXXL	
95	19	EW	177	VLDAGFRL	XLXXXXXXL	
85	17	POL	741	VLSRKYSF	XLXXXXXF	
85	17	POL	741	VLSRKYSFPW	XLXXXXXXXW	
80	16	POL	542	VLGAKSVQHL	XVXXXXXXXL	
85	17	POL	740	VVLSRKYSF	XVXXXXXXXF	
95	19	POL	525	VVRRAPFHCL	XVXXXXXXXL	
95	19	NUC	124	VWIRTPPAY	XWXXXXXXY	
75	15	EW	353	VWLSVIWM	XWXXXXXXM	
90	18	NUC	102	WFHISCLTF	XFXXXXXF	0.0300
95	19	EW	345	WVVLGSPTVW	XFXXXXXXXW	0.0120
95	19	EW	345	WVVLGSPTVWL	XFXXXXXXXL	
80	16	POL	759	WLRGTSF	XIXXXXXF	
80	16	POL	759	WLRGTSFVY	XIXXXXXXXY	
95	19	NUC	125	WIRTPPAY	XIXXXXXY	
80	16	POL	751	WLLGCAANW	XLXXXXXXW	
80	16	POL	751	WLLGCAANWI	XLXXXXXXI	
80	16	POL	751	WLLGCAANWIL	XLXXXXXXXL	
95	19	POL	414	WLSLDVSAAF	XLXXXXXXXF	



## HBV A24 SUPER MOTIF (With binding information)

Conservancy	Freq	Protein	Position	Sequence	String	A*2401
95	19	POL	414	WLSLDVSAIFY	XLXXXXXXXXY	
100	20	ENV	335	WLSLLVPF	XLXXXXXF	
100	20	ENV	335	WLSLLVPFQW	XLXXXXXXXXW	
85	17	NUC	26	WLWGMIDPY	XLXXXXXXXXY	
95	19	ENV	237	WMCLRRFI	XMXXXXXI	
95	19	ENV	237	WMCLRRFI	XMXXXXXI	0.0230
95	19	ENV	237	WMCLRRFIIF	XMXXXXXXF	0.0013
95	19	ENV	237	WMCLRRFIIFL	XMXXXXXXL	
85	17	ENV	359	WMMWYWGPSL	XMXXXXXXL	0.0005
85	17	ENV	359	WMMWYWGPSL	XMXXXXXXY	
100	20	POL	52	WTHKVGNF	XTXXXXXF	
95	19	POL	52	WTHKVGNTGL	XTXXXXXXL	
95	19	ENV	188	WWTSLNFL	XWXXXXL	
85	17	ENV	362	WYWGPSLY	XYXXXXY	0.0001
100	20	POL	147	YLHTLWKAGI	XLXXXXXXXXI	
100	20	POL	147	YLHTLWKAGIL	XLXXXXXXXXL	
100	20	POL	122	YLPDKGI	XLXXXXXI	
100	20	POL	122	YLPDKGIKPY	XLXXXXXXXXY	
90	18	NUC	118	YLVSGVW	XLXXXXW	
90	18	NUC	118	YLVSGVWI	XLXXXXXI	
85	17	POL	746	YTSFPWLL	XTXXXXL	

411

HBV A24 SUPER MOTIF (With binding information)

TABLE XI

## HBV B07 SUPER MOTIF (with binding information)

Conservancy	Frequency	Protein	Position	Sequence	P2	C-term	AA	B*0702	B*3501 CIR	B*5101	B*5301	B*5401
75	15	X	146	APCNFTSA	P	A	9					
95	19	POL	633	APFIOQSY	P	Y	8	0.0001	0.0012	0.0019	0.0002	0.0002
95	19	POL	633	APFIOQGYPA	P	A	10	0.0029	0.0001		0.0002	1.4000
95	19	POL	633	APFIOQGYPAL	P	L	11	0.2300	0.0010	0.0004	-0.0003	0.0093
100	20	BN	232	CPGYRMQL	P	L	9					
80	16	NLC	14	CPTVQASKL	P	L	9					
80	16	NLC	14	CPTVQASKL	P	L	11					
80	16	X	10	DPADYLCL	P	L	9					
80	16	BN	122	DPHRLY	P	Y	8					
90	18	POL	778	DPKRGRL	P	L	9	0.0120	0.0001	0.0001	0.0001	0.0001
90	18	NLC	33	DPYKEFGA	P	A	8	0.0001	0.0001	0.0019	0.0002	0.0019
75	15	BN	130	FPAGGSSGTV	P	V	11					
90	18	BN	14	FPDQLDPA	P	A	9	0.0002	0.0016	0.0003	0.0011	0.0021
85	17	BN	14	FPDQLDPAF	P	F	10	0.0001	0.5250	0.0665	0.5400	0.0199
95	19	POL	530	FPHCLAFSY	P	Y	9	0.0001	0.2200	0.0900	0.0790	0.0480
95	19	POL	530	FPHCLAFSYM	P	M	10					
75	15	POL	749	FPMLGCA	P	A	8					
75	15	POL	749	FPMLGCAA	P	A	9					
75	15	POL	749	FPMLGCAANV	P	W	11	0.0900	0.0001	0.0001	0.0002	0.0035
90	18	X	67	GPCALRTSA	P	A	10	0.0001	0.0001	0.0002	0.0001	0.0002
95	19	POL	19	GRLEELPRL	P	L	10	0.0002	0.0001	0.0001	-0.0003	0.0001
90	18	POL	18	GRLEELPRLA	P	A	8	-0.0001	0.0001	0.0110	0.0002	0.0065
95	19	BN	173	GRLYMDGF	P	F	10	0.0003	0.0001	0.0001	0.0002	0.0008
95	19	BN	173	GRLYMDGF	P	F	11	0.0011	0.0001	0.0001	0.0001	0.0009
85	17	POL	97	GRLYMDGF	P	L	11	0.0001	0.0001	0.0001	-0.0003	0.0001
100	20	POL	429	HPAAMPIL	P	L	8	0.0650	0.0004	0.3100	0.0037	0.0001
100	20	POL	429	HPAAMPIL	P	L	9	0.0980	0.0270	0.0110	0.0500	0.0120
85	17	POL	429	HPAAMPIL	P	V	10	0.0160	0.0020	0.0078	0.0140	0.0170
80	16	POL	495	HPMLGFRH	P	I	10					
100	20	BN	313	IPPSSWA	P	A	8	0.0004	0.0004	0.0019	0.0002	0.0600
100	20	BN	313	IPPSSWAF	P	F	9	0.1300	2.7679	2.3500	0.7450	0.0034
80	16	BN	313	IPPSSWAF	P	A	10	0.0013	0.0024		0.0014	0.4500
80	16	POL	504	IPMGVGLSPFL	P	F	11					
80	16	POL	504	IPMGVGLSPFL	P	L	11					
90	18	BN	191	IPQSLDSW	P	W	8					
90	18	BN	191	IPQSLDSW	P	W	9					
80	16	BN	315	IPSSWAF	P	A	8					
100	20	POL	50	IPWTHKGNF	P	F	10	0.0013	0.0001	0.0007	0.0001	0.0002
100	20	BN	379	LPFTQLW	P	W	8	0.0001	0.0001	0.0360	0.1400	0.0035
100	20	BN	379	LPFTQLW	P	V	9	0.0002	0.0079	0.0002	0.0006	0.0002
100	20	BN	379	LPFTQLW	P	Y	10	0.0002	0.0001	0.0043	0.0139	0.0021
100	20	BN	379	LPFTQLW	P	I	11	0.0002	0.0001	0.0002	0.0002	0.0002
85	17	POL	712	LPHTAEL	P	L	8					
85	17	POL	712	LPHTAEL	P	L	9	0.0040	0.0630	0.0052	0.3100	0.0005
85	17	POL	712	LPHTAEL	P	A	10	0.0018	0.0011	0.0016	0.0016	0.3300
85	17	POL	712	LPHTAEL	P	A	11	0.0090	0.0027	-0.0003	0.0120	2.7500
80	16	X	89	LPKVLKRTL	P	L	10					
100	20	POL	123	LPDKGKPY	P	Y	10	0.0001	0.0290	0.0002	0.0003	0.0002
100	20	POL	123	LPDKGKPY	P	Y	11	-0.0002	0.0009	0.0001	0.0007	0.0001
95	19	X	58	LPVCAFSSA	P	A	9	0.0480	0.0710	0.0110	0.0009	19.0000

0.0001 0.0002 0.0003 0.0004 0.0005 0.0006 0.0007 0.0008 0.0009 0.0010 0.0011 0.0012 0.0013 0.0014 0.0015 0.0016 0.0017 0.0018 0.0019 0.0020 0.0021 0.0022 0.0023 0.0024 0.0025 0.0026 0.0027 0.0028 0.0029 0.0030 0.0031 0.0032 0.0033 0.0034 0.0035 0.0036 0.0037 0.0038 0.0039 0.0040 0.0041 0.0042 0.0043 0.0044 0.0045 0.0046 0.0047 0.0048 0.0049 0.0050 0.0051 0.0052 0.0053 0.0054 0.0055 0.0056 0.0057 0.0058 0.0059 0.0060 0.0061 0.0062 0.0063 0.0064 0.0065 0.0066 0.0067 0.0068 0.0069 0.0070 0.0071 0.0072 0.0073 0.0074 0.0075 0.0076 0.0077 0.0078 0.0079 0.0080 0.0081 0.0082 0.0083 0.0084 0.0085 0.0086 0.0087 0.0088 0.0089 0.0090 0.0091 0.0092 0.0093 0.0094 0.0095 0.0096 0.0097 0.0098 0.0099 0.0100 0.0101 0.0102 0.0103 0.0104 0.0105 0.0106 0.0107 0.0108 0.0109 0.0110 0.0111 0.0112 0.0113 0.0114 0.0115 0.0116 0.0117 0.0118 0.0119 0.0120 0.0121 0.0122 0.0123 0.0124 0.0125 0.0126 0.0127 0.0128 0.0129 0.0130 0.0131 0.0132 0.0133 0.0134 0.0135 0.0136 0.0137 0.0138 0.0139 0.0140 0.0141 0.0142 0.0143 0.0144 0.0145 0.0146 0.0147 0.0148 0.0149 0.0150 0.0151 0.0152 0.0153 0.0154 0.0155 0.0156 0.0157 0.0158 0.0159 0.0160 0.0161 0.0162 0.0163 0.0164 0.0165 0.0166 0.0167 0.0168 0.0169 0.0170 0.0171 0.0172 0.0173 0.0174 0.0175 0.0176 0.0177 0.0178 0.0179 0.0180 0.0181 0.0182 0.0183 0.0184 0.0185 0.0186 0.0187 0.0188 0.0189 0.0190 0.0191 0.0192 0.0193 0.0194 0.0195 0.0196 0.0197 0.0198 0.0199 0.0200 0.0201 0.0202 0.0203 0.0204 0.0205 0.0206 0.0207 0.0208 0.0209 0.0210 0.0211 0.0212 0.0213 0.0214 0.0215 0.0216 0.0217 0.0218 0.0219 0.0220 0.0221 0.0222 0.0223 0.0224 0.0225 0.0226 0.0227 0.0228 0.0229 0.0230 0.0231 0.0232 0.0233 0.0234 0.0235 0.0236 0.0237 0.0238 0.0239 0.0240 0.0241 0.0242 0.0243 0.0244 0.0245 0.0246 0.0247 0.0248 0.0249 0.0250 0.0251 0.0252 0.0253 0.0254 0.0255 0.0256 0.0257 0.0258 0.0259 0.0260 0.0261 0.0262 0.0263 0.0264 0.0265 0.0266 0.0267 0.0268 0.0269 0.0270 0.0271 0.0272 0.0273 0.0274 0.0275 0.0276 0.0277 0.0278 0.0279 0.0280 0.0281 0.0282 0.0283 0.0284 0.0285 0.0286 0.0287 0.0288 0.0289 0.0290 0.0291 0.0292 0.0293 0.0294 0.0295 0.0296 0.0297 0.0298 0.0299 0.0300 0.0301 0.0302 0.0303 0.0304 0.0305 0.0306 0.0307 0.0308 0.0309 0.0310 0.0311 0.0312 0.0313 0.0314 0.0315 0.0316 0.0317 0.0318 0.0319 0.0320 0.0321 0.0322 0.0323 0.0324 0.0325 0.0326 0.0327 0.0328 0.0329 0.0330 0.0331 0.0332 0.0333 0.0334 0.0335 0.0336 0.0337 0.0338 0.0339 0.0340 0.0341 0.0342 0.0343 0.0344 0.0345 0.0346 0.0347 0.0348 0.0349 0.0350 0.0351 0.0352 0.0353 0.0354 0.0355 0.0356 0.0357 0.0358 0.0359 0.0360 0.0361 0.0362 0.0363 0.0364 0.0365 0.0366 0.0367 0.0368 0.0369 0.0370 0.0371 0.0372 0.0373 0.0374 0.0375 0.0376 0.0377 0.0378 0.0379 0.0380 0.0381 0.0382 0.0383 0.0384 0.0385 0.0386 0.0387 0.0388 0.0389 0.0390 0.0391 0.0392 0.0393 0.0394 0.0395 0.0396 0.0397 0.0398 0.0399 0.0400 0.0401 0.0402 0.0403 0.0404 0.0405 0.0406 0.0407 0.0408 0.0409 0.0410 0.0411 0.0412 0.0413 0.0414 0.0415 0.0416 0.0417 0.0418 0.0419 0.0420 0.0421 0.0422 0.0423 0.0424 0.0425 0.0426 0.0427 0.0428 0.0429 0.0430 0.0431 0.0432 0.0433 0.0434 0.0435 0.0436 0.0437 0.0438 0.0439 0.0440 0.0441 0.0442 0.0443 0.0444 0.0445 0.0446 0.0447 0.0448 0.0449 0.0450 0.0451 0.0452 0.0453 0.0454 0.0455 0.0456 0.0457 0.0458 0.0459 0.0460 0.0461 0.0462 0.0463 0.0464 0.0465 0.0466 0.0467 0.0468 0.0469 0.0470 0.0471 0.0472 0.0473 0.0474 0.0475 0.0476 0.0477 0.0478 0.0479 0.0480 0.0481 0.0482 0.0483 0.0484 0.0485 0.0486 0.0487 0.0488 0.0489 0.0490 0.0491 0.0492 0.0493 0.0494 0.0495 0.0496 0.0497 0.0498 0.0499 0.0500 0.0501 0.0502 0.0503 0.0504 0.0505 0.0506 0.0507 0.0508 0.0509 0.0510 0.0511 0.0512 0.0513 0.0514 0.0515 0.0516 0.0517 0.0518 0.0519 0.0520 0.0521 0.0522 0.0523 0.0524 0.0525 0.0526 0.0527 0.0528 0.0529 0.0530 0.0531 0.0532 0.0533 0.0534 0.0535 0.0536 0.0537 0.0538 0.0539 0.0540 0.0541 0.0542 0.0543 0.0544 0.0545 0.0546 0.0547 0.0548 0.0549 0.0550 0.0551 0.0552 0.0553 0.0554 0.0555 0.0556 0.0557 0.0558 0.0559 0.0560 0.0561 0.0562 0.0563 0.0564 0.0565 0.0566 0.0567 0.0568 0.0569 0.0570 0.0571 0.0572 0.0573 0.0574 0.0575 0.0576 0.0577 0.0578 0.0579 0.0580 0.0581 0.0582 0.0583 0.0584 0.0585 0.0586 0.0587 0.0588 0.0589 0.0590 0.0591 0.0592 0.0593 0.0594 0.0595 0.0596 0.0597 0.0598 0.0599 0.0600 0.0601 0.0602 0.0603 0.0604 0.0605 0.0606 0.0607 0.0608 0.0609 0.0610 0.0611 0.0612 0.0613 0.0614 0.0615 0.0616 0.0617 0.0618 0.0619 0.0620 0.0621 0.0622 0.0623 0.0624 0.0625 0.0626 0.0627 0.0628 0.0629 0.0630 0.0631 0.0632 0.0633 0.0634 0.0635 0.0636 0.0637 0.0638 0.0639 0.0640 0.0641 0.0642 0.0643 0.0644 0.0645 0.0646 0.0647 0.0648 0.0649 0.0650 0.0651 0.0652 0.0653 0.0654 0.0655 0.0656 0.0657 0.0658 0.0659 0.0660 0.0661 0.0662 0.0663 0.0664 0.0665 0.0666 0.0667 0.0668 0.0669 0.0670 0.0671 0.0672 0.0673 0.0674 0.0675 0.0676 0.0677 0.0678 0.0679 0.0680 0.0681 0.0682 0.0683 0.0684 0.0685 0.0686 0.0687 0.0688 0.0689 0.0690 0.0691 0.0692 0.0693 0.0694 0.0695 0.0696 0.0697 0.0698 0.0699 0.0700 0.0701 0.0702 0.0703 0.0704 0.0705 0.0706 0.0707 0.0708 0.0709 0.0710 0.0711 0.0712 0.0713 0.0714 0.0715 0.0716 0.0717 0.0718 0.0719 0.0720 0.0721 0.0722 0.0723 0.0724 0.0725 0.0726 0.0727 0.0728 0.0729 0.0730 0.0731 0.0732 0.0733 0.0734 0.0735 0.0736 0.0737 0.0738 0.0739 0.0740 0.0741 0.0742 0.0743 0.0744 0.0745 0.0746 0.0747 0.0748 0.0749 0.0750 0.0751 0.0752 0.0753 0.0754 0.0755 0.0756 0.0757 0.0758 0.0759 0.0760 0.0761 0.0762 0.0763 0.0764 0.0765 0.0766 0.0767 0.0768 0.0769 0.0770 0.0771 0.0772 0.0773 0.0774 0.0775 0.0776 0.0777 0.0778 0.0779 0.0780 0.0781 0.0782 0.0783 0.0784 0.0785 0.0786 0.0787 0.0788 0.0789 0.0790 0.0791 0.0792 0.0793 0.0794 0.0795 0.0796 0.0797 0.0798 0.0799 0.0800 0.0801 0.0802 0.0803 0.0804 0.0805 0.0806 0.0807 0.0808 0.0809 0.0810 0.0811 0.0812 0.0813 0.0814 0.0815 0.0816 0.0817 0.0818 0.0819 0.0820 0.0821 0.0822 0.0823 0.0824 0.0825 0.0826 0.0827 0.0828 0.0829 0.0830 0.0831 0.0832 0.0833 0.0834 0.0835 0.0836 0.0837 0.0838 0.0839 0.0840 0.0841 0.0842 0.0843 0.0844 0.0845 0.0846 0.0847 0.0848 0.0849 0.0850 0.0851 0.0852 0.0853 0.0854 0.0855 0.0856 0.0857 0.0858 0.0859 0.0860 0.0861 0.0862 0.0863 0.0864 0.0865 0.0866 0.0867 0.0868 0.0869 0.0870 0.0871 0.0872 0.0873 0.0874 0.0875 0.0876 0.0877 0.0878 0.0879 0.0880 0.0881 0.0882 0.0883 0.0884 0.0885 0.0886 0.0887 0.0888 0.0889 0.0890 0.0891 0.0892 0.0893 0.0894 0.0895 0.0896 0.0897 0.0898 0.0899 0.0900 0.0901 0.0902 0.0903 0.0904 0.0905 0.0906 0.0907 0.0908 0.0909 0.0910 0.0911 0.0912 0.0913 0.0914 0.0915 0.0916 0.0917 0.0918 0.0919 0.0920 0.0921 0.0922 0.0923 0.0924 0.0925 0.0926 0.0927 0.0928 0.0929 0.0930 0.0931 0.0932 0.0933 0.0934 0.0935 0.0936 0.0937 0.0938 0.0939 0.0940 0.0941 0.0942 0.0943 0.0944 0.0945 0.0946 0.0947 0.0948 0.0949 0.0950 0.0951 0.0952 0.0953 0.0954 0.0955 0.0956 0.0957 0.0958 0.0959 0.0960 0.0961 0.0962 0.0963 0.0964 0.0965 0.0966 0.0967 0.0968 0.0969 0.0970 0.0971 0.0972 0.0973 0.0974 0.0975 0.0976 0.0977 0.0978 0.0979 0.0980 0.0981 0.0982 0.0983 0.0984 0.0985 0.0986 0.0987 0.0988 0.0989 0.0990 0.0991 0.0992 0.0993 0.0994 0.0995 0.0996 0.0997

## HBV B07 SUPER MOTIF (with binding information)

Conservancy	Frequency	Protein	Position	Sequence	P2	C-term	AA	B*0702	B*3501	CIR	B*5101	B*5301	B*5401
80	16	POL	611	LPVNRPIIDW	P	W	9						
80	16	POL	611	LPVNRPIIDW	P	V	11						
80	16	POL	433	MPILVGGSG	P	L	11						
100	20	POL	1	MPISYGHF	P	F	8	0.0001	0.0097		0.0120	0.0370	0.0190
75	15	POL	1	MPISYGHF	P	L	11	0.0120	0.0001		0.0001	-0.0003	0.0001
90	18	POL	774	NPADPSRGFL	P	L	11	0.0012	0.0021		0.0001	0.0028	0.0001
95	19	BNV	9	NRIGFPHQL	P	L	11						
75	15	POL	571	NPRIKTKFW	P	W	8						
75	15	POL	571	NPRIKTKFW	P	Y	10						
95	19	NJC	129	PPAYRPPNA	P	A	9	0.0001	0.0001		0.0001	0.0002	0.0003
95	19	NJC	129	PPAYRPPNA	P	I	11	0.0003	0.0001		0.0001	-0.0003	0.0001
85	17	BNV	58	PPHGGILGW	P	W	9	0.0001	0.0002		0.0001	0.0003	0.0002
100	20	NJC	134	PPNAPILSTL	P	L	10	0.0001	0.0001		0.0035	0.0001	0.0002
80	16	POL	615	RPIDMKVQCF	P	I	11	0.0076	0.0001		0.0280	0.0002	0.0002
100	20	NJC	133	RPPIAPIL	P	L	8	0.1300	0.0001		0.0018	-0.0003	0.0001
100	20	NJC	133	RPPIAPILSTL	P	L	11	-0.0002	0.0001		0.0001	-0.0003	0.0011
100	20	NJC	44	SPEHCSPHITA	P	A	11	0.5600	0.0009		0.0180	0.0009	0.0093
95	19	POL	511	SPFLAOF	P	F	8	0.0820	0.0001		0.0001	-0.0003	12.0500
100	20	NJC	49	SPHHTALROA	P	A	10	0.0012	0.0001		0.0001	0.0002	0.0035
100	20	NJC	49	SPHHTALROA	P	I	11	0.5800	0.0001		0.0004	0.0005	0.0002
85	17	BNV	87	SPKQGL	P	L	8						
85	17	POL	808	SPSVPSHL	P	L	8						
75	15	BNV	350	SPTWMLSV	P	Y	8						
75	15	BNV	350	SPTWMLSV	P	I	9						
75	15	BNV	350	SPTWMLSV	P	W	10						
75	15	BNV	350	SPTWMLSV	P	M	11	0.3900	0.0001		0.0019	0.0002	0.0002
95	19	POL	658	SPTYKARL	P	L	8	0.0078	0.0001		0.0013	0.0001	0.0015
90	18	POL	354	TPARVTGGV	P	V	9	0.3200	0.1000		0.0001	0.0039	0.0006
90	18	POL	354	TPARVTGGV	P	F	10	0.0950	0.0001		0.0001	0.0005	0.0005
90	18	POL	354	TPARVTGGV	P	L	11						
95	19	NJC	128	TPPAYRPPNA	P	A	10	0.0001	0.0001			0.0002	0.0100
75	15	BNV	57	TPHGGIL	P	L	8						
75	15	BNV	57	TPHGGIL	P	W	10						
80	16	POL	691	TPHGGIL	P	A	8						
75	15	POL	691	TPHGGIL	P	I	9						
75	15	BNV	340	TPHGGIL	P	V	8	0.0010	0.0001		0.0100	0.0002	0.1100
95	19	BNV	340	TPHGGIL	P	L	10	0.0011	0.0001		0.0100	0.0001	0.0025
95	19	BNV	340	TPHGGIL	P	L	10	0.0006	0.0001		0.0004	0.0001	0.0002
95	19	POL	398	VPNLSQSLNL	P	L	11	0.0004	0.0001		0.0001	-0.0003	0.0002
90	18	POL	769	VPNLSQSLNL	P	L	8	0.0011	0.0001		0.0070	0.0002	1.0000
95	19	POL	393	WPKGAVPNL	P	L	9	0.0054	0.0002		0.0015	0.0001	0.0015
95	19	POL	640	YPALMPLY	P	Y	8	0.0004	0.2600		0.0340	0.0450	0.0056
95	19	POL	640	YPALMPLY	P	A	9	0.0180	0.0400		0.0340	0.0140	16.0000
95	19	POL	640	YPALMPLY	P	A	11	0.0040	0.0001		0.0470	0.0320	0.0700

96

HBV B07 SUPER MOTIF (with binding information)



# HBV B27 Super Motif

Protein	Sequence	Position in HBV	No of Amino Acids	Sequence Frequency	Conservancy (%)
AYW	LHTLWKAGI	148	9	20	100
AYR	LKLMPARF	107	9	15	75
AYW	LRGLPVCAF	55	9	19	95
AYW	LRGTSPPV	761	9	16	80
AYW	LRRIIFLF	240	9	15	75
AYW	PHCLAFSYM	531	9	19	95
AYW	PHHTALROA	50	9	20	100
AYW	PKVLHKRTL	90	9	17	85
AYR	QHFRKLTL	6	9	15	75
AYW	QRWGILGF	623	9	18	90
AYW	RKIPMGVGL	502	9	16	80
AYW	RKLPVNIPI	609	9	16	80
AYW	RKYSFPWL	744	9	17	85
AYW	RRAFPICLA	527	9	19	95
AYW	RRIIFLFI	241	9	15	75
AYR	RRLKLMPA	105	9	15	75
AYW	RRAVEDJNL	35	9	18	90
AYW	SKLQIGMLW	20	9	17	85
AYW	SRKYTSPPW	743	9	17	85
AYW	TRHYLTLW	144	9	20	100
AYW	VHFASPLHV	819	9	16	80
AYW	VRFSLSL	331	9	16	80
AYW	VRRAPPHCL	526	9	19	95
AYW	YRPPNAPIL	132	9	20	100
AYW	YRWMLRRF	235	9	19	95
AYW	AHLSLRGLPV	51	10	18	90
AYW	AKSVQHLESL	546	10	17	85
AYW	ARDVLCRPV	12	10	15	75
AYW	ARVTGVFLV	356	10	18	90
AYW	EHCSPHHTAL	46	10	20	100
AYW	FRIKPMGVGL	501	10	16	80
AYW	FRIKLPVNIPI	608	10	16	80
AYR	GRETVELEV	140	10	15	75
AYW	HHITALROAIL	51	10	19	95
AYW	HKVGNTGLY	54	10	19	95
AYW	KRWGYSLNFM	576	10	17	85
AYW	LHLVSHPIIL	490	10	16	80
AYW	LHPAAMPILL	428	10	20	100
AYW	LHTLWKAGIL	148	10	20	100
AYR	LKLMPARFY	107	10	15	75
AYW	LRRIIFLFI	240	10	15	75
AYW	NKTKRWGYSL	573	10	15	75
AYW	NRRVAEDJNL	34	10	17	85
AYW	PHHTALROAI	50	10	20	100
AYW	PHLLVGSSGL	434	10	16	80

# HBV B27 Super Motif

Protein	Sequence	Position in HBV	No. of Amino Acids	Sequence Frequency	Conservancy (%)
AYW	QRVGLLGFA	623	10	18	90
AYW	RHYLHTLWKA	145	10	20	100
AYW	RKYSFPWL	744	10	17	85
AYW	RRAEPHCLAF	527	10	19	95
AYW	RRIIIFLIL	241	10	15	75
AYW	SRKYSFPWL	743	10	17	85
AYW	SRLVDFSQF	375	10	19	95
AYW	THKVGNIETGL	53	10	19	95
AYW	TKRWGYSLNF	575	10	17	85
AYW	TKYLPDKGI	120	10	20	100
AYW	TRILTIPOS	186	10	16	80
AYW	VHIFASPLHYA	819	10	16	80
AYW	VRFESMLSLV	331	10	16	80
AYW	VIRAEPHCLA	526	10	19	95
AYW	WKKQQRIVGL	619	10	17	85
AYW	YFWMCLRRFI	235	10	19	95
AYW	DHGAILSLRGL	48	11	19	95
AYW	IHLNPNKTKRW	568	11	15	75
AYW	IHTAELLACF	714	11	17	85
AYW	LHPAAMPILLY	428	11	17	85
AYW	LHTLWKAGILY	148	11	20	100
AYW	LROALCWGEL	55	11	18	90
AYW	LRIEIIIFLIL	240	11	15	75
AYW	PHHTALRQAIL	50	11	19	95
AYW	PKFAVPNLQSL	394	11	19	95
AYW	PKVLHKRTLGL	90	11	17	85
AYW	PRTPARVITGGV	352	11	18	90
AYW	QRIVGLLGFAA	623	11	18	90
AYW	RKLPVNIPIIDW	609	11	16	80
AYW	RRIIIFLIL	241	11	15	75
AYW	RRLKIMPARF	105	11	15	75
AYW	SHPILGFRI	494	11	16	80
AYW	SKLQGLWLGGM	20	11	17	85
AYW	SRKYSFPWL	743	11	17	85
AYW	THKVGNIETGLY	53	11	19	95
AYW	TKRWGYSLNFM	575	11	17	85
AYW	TRHYLHTLWKA	144	11	20	100
AYW	VHIFASPLHYAW	819	11	16	80
AYW	VIRAEPHCLAF	526	11	19	95
AYW	WKKQQRIVGL	619	11	17	85
AYW	YFWMCLRRFI	235	11	19	95

TABLE XIII  
HBV B58 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
POL	AAMPILLY	431	8	17	85
NUC	ASALYREA	34	8	17	85
POL	ASFCGSPY	166	8	20	100
NUC	ASKLCLGW	19	8	18	90
POL	ASPLHVAV	822	8	16	80
BN	ASVRFSWL	329	8	16	80
POL	ATPTGWGL	690	8	19	95
X	CALRFTSA	69	8	18	90
NUC	CSPHHTAL	48	8	20	100
POL	CSVVRPAF	523	8	19	95
POL	ESRLWDF	374	8	19	95
NUC	ETVLELY	142	8	15	75
POL	FARSRGA	724	8	17	85
POL	FASPLHYA	821	8	16	80
POL	FSPTYKAF	821	8	19	95
X	FSSAGPCA	658	8	19	95
BN	FSMLSLLY	63	8	20	100
POL	FSYMDDV	333	8	18	90
POL	FTOCGYPA	536	8	19	95
POL	FTSAICSV	635	8	19	95
POL	GAKSVQHL	518	8	17	85
POL	GTDNSWL	545	8	18	90
POL	HTAELLAA	735	8	17	85
NUC	HTALHQA	715	8	20	100
POL	HTLWKAGI	52	8	20	100
POL	LAOFTSAI	149	8	19	95
NUC	LSLPSDF	515	8	19	95
POL	LSLDVSAA	45	8	19	95
BN	LSLLVPFV	415	8	20	100
X	LSLRGLPV	336	8	19	95
POL	LSRKYTSF	53	8	17	85
POL	LSNLSQL	742	8	18	90
POL	LSMLSVDV	408	8	20	100
NUC	LTFGRETV	412	8	19	95
X	MSTTDLEA	108	8	16	80
NUC	NAPILSTL	103	8	20	100
POL	PAAMPPLL	136	8	20	100
POL	PALMPLYA	430	8	19	95
X	PARDVLC	641	8	16	80
POL	PARVTGGV	11	8	18	90
NUC	PAYRPPNA	355	8	19	95
POL	PSRGLQL	130	8	18	90
POL	PTGWGLAI	779	8	15	75
POL	PTTGRTSL	692	8	17	85
NUC	PTVOASKL	797	8	16	80
NUC		15	8		

# HBV B58 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
BW	PTWLSLVI	351	8	15	75
POL	RAFPHCLA	528	8	19	95
X	RTLGLSAM	96	8	24	120
NJC	SALYREAL	35	8	18	90
X	SSAGPCAL	64	8	19	95
BW	SSGTVNPV	136	8	15	75
BW	SSKPRQGM	5	8	18	90
NJC	STLPETTV	141	8	20	100
X	STTDLEAY	104	8	15	75
NJC	TALRQAIL	53	8	19	95
POL	TSAICSVV	519	8	19	95
BW	TSGRLGPL	168	8	16	80
X	TTDLEAYF	105	8	15	75
POL	TTGRTSLY	798	8	17	85
NJC	VSWPKFAV	391	8	19	95
POL	VSYNNVNM	115	8	20	100
NJC	VTGGVFLV	358	8	20	100
BW	WSPQACGI	66	8	17	85
POL	WTHKVGNF	52	8	20	100
POL	YSLNFMGY	580	8	17	85
POL	YTSFPMLL	746	8	17	85
POL	AAFTQCGY	632	9	19	95
NJC	ASALYREAL	34	9	17	85
NJC	ASKLCLGWL	19	9	18	90
POL	ATPTGWGLA	690	9	16	80
POL	CSRNLVSL	471	9	16	80
POL	DATPTGWGL	689	9	19	95
BW	DSMWTSINF	196	9	19	95
POL	EAGPLEEL	17	9	20	100
POL	FADATPTGW	687	9	19	95
POL	FASPLHNAV	821	9	16	80
POL	FAVPILOS	396	9	19	95
POL	FSPTYKAE	658	9	19	95
X	FSSAGPCAL	63	9	19	95
POL	FSSYMDVVL	536	9	18	90
POL	FTFSPTYKA	656	9	19	95
POL	FTGLYSSTV	59	9	18	90
POL	FTOCGYPAL	635	9	19	95
POL	FTSAICSVV	518	9	19	95
X	GAHLSLRGL	50	9	19	95
NJC	HTALRQAIL	52	9	19	95
POL	HTLWKAGIL	149	9	20	100
POL	KSVQHLESL	547	9	17	85
POL	KTKRMGYSL	574	9	19	95
POL	LAFSYMDV	534	9	18	90



# HBV B58 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
NJC	LSRLPSDF	45	9	19	95
POL	LSLDVSAF	415	9	19	95
POL	LSPLLAOF	510	9	19	95
ENV	LSPTWMLSV	349	9	15	75
NJC	LSTLPETTV	140	9	20	100
ENV	LSYPANPLGF	16	9	15	75
POL	LSYQHFFKL	3	9	15	75
NJC	LTFGREVL	137	9	15	75
POL	LTNLLSNL	404	9	18	90
POL	LTNKKRL	99	9	17	85
POL	MSTTDLEAY	103	9	15	75
X	NSVLSRKY	738	9	18	90
POL	PAAMPHLLV	430	9	17	85
POL	PARVTGVF	355	9	18	90
POL	PTTGRTSLY	797	9	17	85
ENV	PTWMLSVW	351	9	15	75
POL	QAFTEPTV	654	9	19	95
NJC	QALCWGEL	57	9	18	90
NJC	QASKLCLGW	18	9	16	80
POL	RAEPHCLAF	528	9	19	95
ENV	RTGDPAPNM	167	9	16	80
X	SAGPCALRF	65	9	18	90
POL	SASFCGSPY	165	9	20	100
POL	SSNLSQLSL	409	9	18	90
ENV	SSSGIVNPV	135	9	15	75
NJC	STLPETTV	141	9	20	100
X	STDLEAYF	104	9	15	75
POL	TAEILAACF	716	9	17	85
NJC	TASALYREA	33	9	16	80
POL	TSFVYVPSA	764	9	16	80
ENV	TSGRLGPL	168	9	15	75
POL	TTGRTSLYA	798	9	17	85
POL	VSIPVTHKV	48	9	20	100
ENV	WSPQAGIL	66	9	17	85
ENV	WSSKPRQGM	4	9	18	90
POL	YSHPLLGF	493	9	16	80
POL	YSLNFMGV	580	9	15	75
POL	ASFCGSPYSW	166	10	20	100
NJC	ASKLCLGMLW	19	10	17	85
ENV	ASVRFSLSL	329	10	16	80
POL	ATPTGWGLAI	690	10	15	75
X	CAFSSAGPCA	61	10	19	95
ENV	CTCIPPSW	310	10	20	100
ENV	CTIPAOGTSM	298	10	16	80
POL	DAIPTGWGLA	689	10	16	80

# HBV B58 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
ENV	DSWMTSLNFL	196	10	18	90
NUC	DTASALYREA	32	10	16	80
POL	FAAPTOCGY	631	10	19	85
ENV	FSMLSLVPF	333	10	20	100
POL	FTFSPTYKAF	656	10	19	95
POL	FTOCGYPALM	635	10	38	190
ENV	GSSSGTVNPFV	134	10	15	75
ENV	GTNLSPNPRL	13	10	15	75
POL	GTSFVYVPSA	763	10	16	80
POL	HTAELLAACF	715	10	17	85
POL	HTLWKAGILY	149	10	20	100
POL	LAFSYMDDVY	534	10	18	90
POL	LSLDVSAAFY	415	10	19	95
ENV	LSILVPRQW	336	10	20	100
X	LSLRGLPVCA	53	10	19	95
ENV	LSPTVWLSVI	349	10	15	75
POL	LSRKTSFEPW	742	10	17	85
POL	LSNL.SWLSL	408	10	18	90
NUC	LSLTPETTVV	140	10	20	100
POL	LSWL.SLDVSA	412	10	20	100
POL	LSYCHFRKLL	3	10	15	75
ENV	LTIPOSLSDW	189	10	18	90
X	MSTTDLEAYF	103	10	15	75
POL	PADDPSSRGRL	775	10	18	90
ENV	PAGSSSGTV	131	10	18	90
POL	PALMPLYACI	641	10	19	95
X	PARCNFTSA	145	10	15	75
POL	PARVTGGVRL	355	10	18	90
NUC	PAYRPPNAPI	130	10	19	95
POL	PTTGRTSLYA	797	10	17	85
NUC	PTYQASKLCL	15	10	16	80
ENV	PTVWL.SVWMM	351	10	30	150
ENV	QAGFLLTRI	179	10	16	80
NUC	QALCWGELM	57	10	36	180
ENV	QAMQWNSTTF	107	10	16	80
NUC	QASKLCL.GWML	18	10	16	80
ENV	QSLDSWWTSL	193	10	18	90
POL	RTPPARVTGGV	353	10	18	90
POL	SAICSVVRRA	520	10	19	95
X	SSAGPCALRF	64	10	18	90
POL	TAELLAACFA	716	10	17	85
NUC	TALRALICW	53	10	19	95
NUC	TASALYREAL	33	10	16	80
POL	TSFPMILGCA	747	10	15	75
POL	TSFVYVPSAL	764	10	16	80

# HBV B58 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
ENV	TSGRLGPLV	168	10	15	75
POL	VAEDNLGNL	37	10	19	95
POL	YSLNFMGYVI	580	10	15	75
POL	AACFARSRSGA	721	11	17	85
POL	AAPFTCCGYPA	632	11	19	95
ENV	ASVRFSLLSLL	329	11	16	80
X	CAFSSAGPCAL	61	11	19	95
X	CALRFTSARRM	69	11	26	130
NUC	CSPHHTALRQA	48	11	20	100
ENV	CTCIPIPSSWA	310	11	20	100
POL	DATPTGWGLAI	689	11	15	75
NUC	DTASALYREAL	32	11	16	80
POL	ESRLVDSQF	374	11	19	95
POL	FADATPTGWGL	687	11	19	95
X	FSSAGPCALRF	63	11	18	90
ENV	FSWLSLVPIV	333	11	20	100
POL	FSYMDDWLGA	536	11	18	90
POL	FTFSPTYKAPL	656	11	19	95
X	GAHLSLRGLPV	50	11	18	90
POL	GAKSYQHLESL	545	11	17	85
POL	GTSFYVPSAL	763	11	16	80
POL	HTAELLAACFA	715	11	17	85
NUC	HTALRQALCW	52	11	19	95
NUC	ISCLTFGRETV	105	11	18	90
POL	KTKRMGYSUNF	574	11	17	85
POL	LAFTSYMDVNL	534	11	18	90
POL	LAQFTSAICSV	515	11	19	95
ENV	LSLLVPRQMF	336	11	20	100
X	LSLRGLPYCAF	53	11	19	95
ENV	LSPTVWLSVW	349	11	15	75
POL	LSRKYTSFPWL	742	11	17	85
POL	LSMWLSLDVSA	412	11	19	95
POL	LSYQHFRKILL	3	11	15	75
NUC	LTFGRETVLEY	137	11	15	75
ENV	LTIPOSIDSWW	189	11	18	90
POL	LTNLSSNLSTW	404	11	18	90
ENV	LTRILLTIPOS	185	11	16	80
X	PARDVLCIRPV	11	11	15	75
POL	PARVTGGVRLV	355	11	18	90
NUC	PAYRPPNAPIL	130	11	19	95
ENV	PTVWLSTVIMMM	351	11	28	140
POL	QAFTFSPTYKA	654	11	19	95
ENV	QAGFLLTRILL	179	11	16	80
NUC	QASKLQGLMW	18	11	15	75
POL	QSLTNLSSNL	402	11	18	90

# HBV B58 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
POL	RAAPHQLAFSY	528	11	19	95
POL	RTPARVTGGVF	353	11	18	90
N/C	RTPPAYRPPNA	127	11	19	95
POL	SAICSVRRRAF	520	11	19	95
POL	SASFQSPYSW	165	11	20	100
POL	SSNLSQLSLDV	409	11	18	90
POL	TSAICSVVRA	519	11	19	95
POL	TSFPWLLGCA	747	11	15	75
ENV	TSGRLGPLVL	168	11	15	75
POL	VSWPKFAVPNL	391	11	19	95
POL	WTHKGNFTGL	52	11	19	95
POL	YTSPWLLGCA	746	11	15	75
	237				

TABLE XIV  
HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
N/C	ALLCWGEI	58	8	18	90
POL	APFTIOGY	633	8	19	95
POL	AVPNLIOSL	397	8	19	95
BW	CIPPPSW	312	8	20	100
N/C	QLGMLWGM	23	8	17	85
BW	CLIFLLVL	253	8	20	100
BW	CLRRHIF	239	8	19	95
BW	COFMGIL	622	8	17	85
POL	DIDPYKEF	31	8	18	90
N/C	DLLDTASA	29	8	17	85
N/C	DPRVRGLY	122	8	16	80
BW	DPYKEFGA	33	8	18	90
N/C	DVLCURPV	14	8	19	95
X	ELGEEEL	122	8	16	80
X	ELLAACFA	718	8	18	90
POL	FIIFLIL	243	8	16	80
BW	FILLCLI	248	8	16	80
BW	FLGPLVL	171	8	15	75
BW	FLVVLDDY	256	8	19	95
POL	FPMLGCA	749	8	15	75
BW	FVGLSPV	346	8	19	95
BW	FVGMFVGL	342	8	18	90
POL	FVYVPSAL	766	8	19	95
POL	GLSPFLA	509	8	20	100
BW	GLSPTVWL	348	8	18	90
BW	GMLPYCPL	265	8	19	95
BW	GPLVLQA	173	8	16	80
POL	GVGLSPFL	507	8	16	80
POL	HLYSHPII	491	8	20	100
POL	HPAAMPPL	429	8	16	80
BW	IIFLILL	244	8	16	80
POL	IILGFRII	497	8	18	90
N/C	IILCWGELM	59	8	20	100
BW	IILLCLIF	249	8	16	80
POL	IILRGTSFV	760	8	19	95
BW	ILTIPOS	188	8	20	100
BW	IPIPPSWA	313	8	18	90
BW	IPOSLSW	191	8	16	80
BW	IPSSWAFA	315	8	18	90
POL	IVGALLGFA	625	8	16	80
POL	KIPMGVGL	503	8	17	85
N/C	KIQGMLW	21	8	15	75
POL	KLIMPAPF	108	8	16	80
POL	KLPVNPPI	610	8	19	95
POL	KVGNFTGL	55	8		

# HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
X	KVLHKRTL	91	8	17	85
BW	LIFELVLL	254	8	20	100
POL	LIMPARFY	109	8	20	100
POL	LIAQFTSA	514	8	19	95
BW	LLQLFL	251	8	20	100
NUC	LIDTASAL	30	8	17	85
BW	LIDYQGM	260	8	19	95
POL	LLGCANW	752	8	16	80
POL	LLGFAPF	628	8	19	95
BW	LLGWSPQA	63	8	17	85
BW	LLQLFL	250	8	20	100
BW	LLPIFFQL	378	8	20	100
POL	LLSLGIHL	563	8	19	95
POL	LLSSNLNW	407	8	18	90
BW	LLTRILTI	184	8	16	80
POL	LLVGSSGL	436	8	16	80
BW	LLVLOAGF	175	8	19	95
BW	LLVPEVQW	338	8	20	100
POL	LMPLYACI	643	8	19	95
BW	LPPIFQLW	379	8	20	100
POL	LPHTAEL	712	8	17	85
BW	LOAGFRL	178	8	19	95
POL	LOSLTNLL	401	8	20	100
BW	LWLOAGFE	176	8	19	95
BW	LWPPVQWF	339	8	20	100
NUC	LVSFGWMI	119	8	18	90
POL	LWDFSQF	377	8	20	100
POL	MPLSYGHF	1	8	20	100
NUC	MOLPHLQL	1	8	15	75
BW	MOWNSTTF	109	8	15	75
POL	NLWNSIPW	45	8	16	80
POL	NLQSLTNL	400	8	19	95
BW	NLSVNPRL	15	8	20	100
POL	NPNKTKRW	571	8	15	75
BW	PIFFQLW	380	8	15	75
POL	PIHTAELL	713	8	20	100
BW	PIPSSWAF	314	8	17	85
BW	POSLSWV	192	8	20	100
X	PVCAFFSSA	59	8	18	90
POL	PVNRPPDW	612	8	19	95
X	QLDPARDV	8	8	17	85
POL	RIVGLGF	624	8	16	80
POL	RLKLMPA	106	8	18	90
NUC	RPPNAPIL	133	8	15	75
NUC	ROLWRFIL	98	8	20	100

HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
POL	RVAEDLNL	36	8	19	95
POL	RVHFA SPL	818	8	16	80
POL	RVTGAVRL	357	8	20	100
POL	SIPWTHKV	49	8	20	100
POL	SLDVSAAF	416	8	19	95
POL	SLNFMGYV	581	8	15	75
POL	SPRLAQF	511	8	19	95
BW	SPOAQGL	67	8	17	85
POL	SPSVPSHL	808	8	17	85
BW	SPTWMLSV	350	8	15	75
POL	SPTYKARL	659	8	19	95
BW	SVPNPLGF	17	8	15	75
POL	SVQHLESL	548	8	17	85
POL	SVLSRKY	739	8	18	90
NUC	TLPETTV	142	8	20	100
POL	TLWKAGIL	150	8	20	100
BW	TPHGGAL	57	8	15	75
POL	TPTGWGLA	691	8	16	80
POL	TOCGYPAL	636	8	19	95
POL	TNIEKRRL	100	8	17	85
BW	TVWLSVIV	352	8	15	75
BW	VLDYOGM	259	8	19	95
BW	VLAQGFRL	177	8	19	95
BW	VPRVQWV	340	8	19	95
POL	VPSALNPA	769	8	18	90
NUC	VOASKLCL	17	8	16	80
POL	VLGAKSV	542	8	18	90
POL	WLRLGTSF	759	8	16	80
NUC	WIRTPPAY	125	8	19	95
POL	WLSLDVSA	414	8	20	100
BW	WLSLLVPF	335	8	20	100
BW	WMCLRRH	237	8	19	95
POL	YLHTLWKA	147	8	20	100
POL	YLPDKGI	122	8	20	100
NUC	YLVSEGW	118	8	18	90
POL	YPALMPLY	640	8	19	95
POL	YQHFRKL	5	8	15	75
POL	AICSVVRRRA	521	9	19	95
NUC	ALCWGELM	58	9	18	90
POL	ALMPLYACI	642	9	19	95
NUC	ALRQALCW	54	9	19	95
BW	AMOWNSSTTF	108	9	16	80
X	AMSTTDLEA	102	9	15	75
X	APCNFTSA	146	9	15	75
BW	CIPFPSSWA	312	9	20	100

# HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
BN	CLPLVL	253	9	20	100
BN	CLRFIFL	239	9	19	95
MUC	CLTFGETV	107	9	18	90
BN	CPGRVMQL	232	9	20	100
MUC	CPTVOASKL	14	9	16	80
X	COLDPARDV	7	9	16	80
MUC	DLDITASAL	29	9	17	85
POL	DNLGNLV	40	9	19	95
X	DPARDVLC	10	9	16	80
POL	DPSRGRLQ	778	9	18	90
POL	DVLGAKSV	541	9	18	90
BN	FIFLFL	243	9	16	80
BN	FILLCLIF	248	9	16	80
BN	FLFILLCL	246	9	16	80
POL	FLLAQFTSA	513	9	19	95
POL	FLSLGIHL	562	9	19	95
BN	FLTLRLTI	183	9	16	80
BN	FPDQLDPA	14	9	18	90
POL	FPHCLAFSY	530	9	19	95
POL	FPWLLGCAA	749	9	15	75
BN	FVGLSPTW	346	9	19	95
POL	GLCOVFADA	682	9	17	85
POL	GLLGEAFPE	627	9	19	95
BN	GLLGWSPQA	62	9	17	85
POL	GVGLSPRL	507	9	16	80
MUC	GVWIRTPPA	123	9	19	95
POL	HLVGSSEL	435	9	16	80
X	HLSLRGLPV	52	9	18	90
POL	HLYSHPIIL	491	9	16	80
POL	HPAAMPFIL	429	9	20	100
BN	IIFLFLLL	244	9	16	80
POL	ILGFRIKPM	498	9	16	80
BN	ILLCLIFL	249	9	20	100
POL	ILRGTSFVY	760	9	16	80
BN	IPIPSSWAF	313	9	20	100
BN	IPOSLSWW	191	9	18	90
POL	IVALLGFAA	625	9	18	90
POL	KLHLYSHPI	489	9	19	95
POL	KLIMPARFY	108	9	15	75
POL	KVCOHNVGL	620	9	17	85
POL	KVGNFTGLY	55	9	19	95
POL	LLAQFTSAI	514	9	19	95
BN	LLCLIFLLY	251	9	20	100
MUC	LLDTASALY	30	9	17	85
POL	LLGCANWI	752	9	16	80



# HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
BW	LLCLIFLL	250	9	20	100
BW	LLPIFFCLW	378	9	20	100
NC	LSFLPSDF	44	9	19	95
POL	LLSNLSWL	407	9	18	90
BW	LLVLQAGF	175	9	19	95
BW	LLVPRQWF	338	9	20	100
NC	LLVPHISCL	100	9	18	90
BW	LPFFCLWV	379	9	20	100
POL	LPHTAELL	712	9	17	85
X	LPVCAESSA	58	9	19	95
POL	LPVNRPIDW	611	9	16	80
BW	LVLDYQGM	258	9	19	95
BW	LVLAQGFLL	176	9	18	90
BW	LVPVQWV	339	9	19	95
BW	MMWVWGPSL	360	9	17	85
POL	NLGILNLSI	42	9	19	95
POL	NLSSNLSW	406	9	18	90
POL	NLOSITNLL	400	9	20	100
POL	NLSWLSLV	411	9	18	80
BW	PIFFCLWV	380	9	20	100
POL	PIHTAELLA	713	9	17	85
POL	PILGFRI	496	9	16	80
BW	PIPSWAF	314	9	16	80
POL	PLDKIKPY	124	9	20	100
POL	PLEELPRL	20	9	19	95
BW	PLPIFFCL	377	9	20	100
BW	PLVLQAGF	174	9	19	95
POL	PLPIHTAEL	711	9	16	80
POL	PMGVLSPF	505	9	16	80
NC	PPAYRPPNA	129	9	19	95
BW	PRKQLGW	58	9	17	85
X	QLDPARDVL	8	9	16	80
BW	RLITIPSL	187	9	16	80
POL	RIVGLGFA	624	9	18	90
POL	RLWDESCF	376	9	19	95
POL	RVTGVFLV	357	9	20	100
BW	SLSWWTSL	194	9	19	95
POL	SLDVSAFY	416	9	19	95
BW	SLLVPRQW	337	9	20	100
POL	SUNMGYVI	581	9	15	75
X	SLRGLPVCA	54	9	19	95
BW	SPTWLSVI	350	9	15	75
BW	SVFSLSL	330	9	16	80
BW	TIPOSLSW	190	9	18	90
POL	TLWKAGILY	150	9	20	100

# HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
POL	TPARVITGV	354	9	18	90
POL	TPTGWGLAI	691	9	15	75
POL	TOCGYPALM	636	9	19	95
NUC	TVQASKLCL	16	9	16	80
BW	TYWLSVIMM	352	9	15	75
X	VLCAPVGA	15	9	19	95
X	VLGGRHKL	133	9	18	90
X	VLHKRTLGL	92	9	17	85
BW	VLDYQGML	259	9	19	95
BW	VLOAGFTLL	177	9	19	95
POL	WLSRKYTSF	741	9	17	85
POL	WILRGTSFV	759	9	16	80
POL	WILGCAANW	751	9	16	80
POL	WLSLDVSAA	414	9	19	95
BW	WLSILVPFV	335	9	20	100
BW	WMQLRRHIL	237	9	19	95
POL	WPKFAVPNL	393	9	19	95
NUC	YLVSEGVMI	118	9	18	90
POL	YMDDVILGA	538	9	19	95
POL	YPALMPLYA	640	9	15	75
POL	YQFFRKILL	5	9	15	75
POL	YVPSALNPA	768	9	18	90
POL	AICSVVRRAF	521	10	19	95
POL	APFTQCGYPA	633	10	19	95
POL	AOFTSAICSV	516	10	19	95
BW	CIPIPSSWAF	312	10	20	100
POL	CLAFSYMDDV	533	10	18	90
NUC	CLGMLWGMDI	23	10	17	85
BW	CLRRFIILF	239	10	15	75
X	COLDPARDVL	7	10	16	80
POL	COIRNGILGF	622	10	17	85
NUC	DIDPYKEFGA	31	10	18	90
NUC	DLLDTASALY	29	10	17	85
X	DVLCAPVGA	14	10	19	95
NUC	ELSFLPSDF	43	10	19	95
BW	FIIPFLILL	243	10	16	80
BW	FILLCLIFL	248	10	16	80
BW	FLFILLCLL	246	10	16	80
BW	FLGPLWLQA	171	10	15	75
POL	FLLAQFTSAI	513	10	19	95
BW	FPDHQLDPAF	14	10	17	85
POL	FPHQLAFSYM	530	10	19	95
BW	FVGLSPTVWL	346	10	19	95
X	PLGGGRHKL	132	10	18	90
X	GLPVCAFFSSA	57	10	19	95

# HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
POL	GLSPFLAOF	509	10	19	95
BW	GLSPTWMLSV	348	10	15	75
NUC	GMDIDPYKEF	29	10	17	85
X	GPCALRTSA	67	10	18	90
POL	GRLEEDPRL	19	10	19	95
BW	GRLYLOAGF	173	10	19	95
POL	GVGLSPFLA	507	10	16	80
NUC	GWIRTPPAY	123	10	19	95
POL	HLNPKTKRW	569	10	15	75
POL	HPAAMPPLLY	429	10	17	85
POL	HPILLGFRIKI	495	10	16	80
POL	ILLGFRIKIPM	497	10	16	80
BW	ILLCLIFLL	249	10	20	100
POL	ILRGTSFVYV	760	10	16	80
NUC	ILSTLPETTV	139	10	20	100
BW	IPIPSWAF	313	10	16	80
POL	IPMGVGLSPF	504	10	16	80
NUC	IPWTHKVGNF	50	10	20	100
POL	KLCGLMWGM	21	10	17	85
POL	KLHLVSHPII	489	10	16	80
POL	KLPVNRPIDW	610	10	16	80
POL	KOAFTESPTY	653	10	19	95
POL	KVQQRWGLL	620	10	17	85
X	KVLHKRTLGL	91	10	17	85
BW	LIFLLVLDY	254	10	19	95
BW	LLCLFLVYL	251	10	20	100
BW	LLDYOGMLPV	260	10	18	90
POL	LLGCANWIL	752	10	16	80
BW	LLCLIFLLV	250	10	20	100
BW	LLPIFFCLWV	378	10	20	100
NUC	LLSLRPSDEF	44	10	19	95
BW	LLVLDYOGM	257	10	19	95
BW	LLVLOAGFEL	175	10	18	90
BW	LLVPEVQWFV	338	10	19	95
BW	LPFFCLWVY	379	10	20	100
POL	LPHTAELLA	712	10	17	85
X	LPKVLHKRTL	89	10	16	80
POL	LPIDGKIPY	123	10	20	100
BW	LVLIDYOGML	258	10	19	95
BW	LVLQAGFLL	176	10	18	90
BW	MMWVWGPPLY	360	10	17	85
POL	NLLSNLSWL	406	10	18	90
BW	NLSVNPPLGF	15	10	15	75
POL	NPNKTKRWGY	571	10	15	75
POL	NVSPWTHIKV	47	10	20	100

HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
POL	PIDMKVCCRI	616	10	17	85
BW	PIFFCLWYI	380	10	20	100
POL	PIHTAELLAA	713	10	17	85
POL	PLDKGKPYV	124	10	20	100
POL	PLEELPILA	20	10	18	90
BW	PLGFFPHQL	10	10	19	95
POL	PLHPAAMPIL	427	10	20	100
BW	PLPIFFCLW	377	10	20	100
POL	PLVLAQGF	174	10	19	95
POL	PLPHIAELL	711	10	16	80
POL	PLSYOHFRKL	2	10	15	75
POL	PLTVNEKRRL	98	10	17	85
POL	PMGVGLSPFL	505	10	16	80
NUC	PPNAPILSTL	134	10	20	100
POL	PVNRPIDWKV	612	10	17	85
NUC	QLLWHHSCL	99	10	18	90
POL	RIVGLIGFAA	624	10	18	90
POL	RLKLMPARF	106	10	15	75
NUC	ROALCWGEL	56	10	18	90
POL	RVHFA SPLHV	818	10	15	75
BW	SLVPPVQWF	337	10	20	100
X	SLRGLPYCAF	54	10	19	95
POL	SLTNLLSSNL	403	10	18	90
NUC	SPHHTALROA	49	10	20	100
BW	SPTVWLSVIW	350	10	15	75
BW	SVRFSLSL	330	10	16	80
BW	TIPOSLSWV	190	10	18	90
POL	TPARVTGGVF	354	10	18	90
NUC	TPPAYRPPNA	128	10	19	95
BW	TPPHGILGW	57	10	15	75
POL	VLGAKSVCHL	543	10	17	85
X	VLGGRHKLY	133	10	18	90
BW	VPPVQMPVQL	340	10	19	95
POL	VPNLQSLTNL	398	10	19	95
NUC	VOASKLCLGW	17	10	16	80
POL	VVLSRYKTSF	740	10	17	85
POL	VVRRAPPHCL	525	10	19	95
POL	WILRGTSFVY	759	10	16	80
POL	WILGCAANMI	751	10	19	95
NUC	WLSLIVSAAF	414	10	16	80
BW	WLWGMIDIPY	26	10	17	85
BW	WMQLRRFIIF	237	10	19	95
BW	WMMWYWGPSL	359	10	17	85
POL	YLIHTLWKAGI	147	10	20	100
BW	YQGMLPVCPIL	263	10	18	90

## HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
POL	YQHFRKLLT	5	10	15	75
POL	APFTQCGYPAL	633	11	19	95
POL	AOFTSAICSV	516	11	19	95
POL	AVPNLQSLTNL	397	11	19	95
BW	CIPIPSSWAF	312	11	16	80
POL	CLAFSYMDDV	533	11	18	90
BW	CLIFLLVLDY	253	11	19	95
BW	CLRRRIELI	239	11	15	75
NC	CPTVQASKLCL	14	11	16	80
POL	COARVGLGFA	622	11	17	85
POL	DNLGNLWVI	40	11	19	95
NC	ELSLPDDF	43	11	19	95
BW	FILLCLIFL	248	11	-16	80
BW	FLHLLCLIF	246	11	16	80
BW	FLWLDYQGM	256	11	19	95
BW	FPAGSSSGIV	130	11	15	75
POL	FPWLGCANW	749	11	15	75
X	FVLGGRHKLV	132	11	18	90
POL	FVYVPSALNPA	766	11	18	90
BW	GLSPTVWLSVI	348	11	15	75
POL	GRLEELPRIA	19	11	18	90
BW	GPLTWMQAGF	173	11	19	95
POL	GPLTWMKRR	97	11	17	85
POL	HLSLRGLPVCA	52	11	18	90
X	HLVSHPIIGF	491	11	16	80
POL	IIFLFIILLCL	244	11	16	80
BW	ILLCLIFLTV	249	11	20	100
NC	ILSTLPETTV	139	11	20	100
BW	ILTIPOSLSW	188	11	18	90
POL	IPMGVGLSPFL	504	11	16	80
POL	IVGLIGFAPE	625	11	18	90
POL	KIPMGVGLSPF	503	11	16	80
POL	KULHYSPIL	489	11	16	80
BW	LLCLIFLVL	251	11	20	100
BW	LLGWSPOAGI	63	11	15	75
BW	LLCLIFLVL	250	11	20	100
BW	LLPIFFQLWY	378	11	20	100
POL	LLSSNLSWLSL	407	11	18	90
BW	LLVLDYQGM	257	11	19	95
BW	LLVLDYQGM	175	11	18	90
NC	LLVLDYQGM	100	11	17	85
BW	LLVLDYQGM	100	11	20	100
POL	LPIFFQLWY	379	11	20	100
POL	LPIHTAELAA	712	11	17	85
POL	LPLDGKIPY	123	11	20	100
POL	LPLVNRPIDWKV	611	11	16	80

# HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
BW	LOAGFLLTRI	178	11	16	80
BW	LVPVQWVGL	339	11	19	95
POL	MPHLYGSSGL	433	11	16	80
POL	MPLSYCHFRKL	1	11	15	75
POL	NLGNLVNIPW	42	11	19	95
POL	NLSMLSLDVSA	411	11	18	90
POL	NPADDPNRGRL	774	11	18	90
BW	NPLGFFPDHQL	9	11	19	95
POL	PIDWAKVQDAV	616	11	17	85
POL	PILGFRKIPM	496	11	16	80
NUC	PILSTLPETTV	138	11	20	100
POL	PLHPAAMPHLL	427	11	20	100
BW	PULPFECQW	377	11	20	100
BW	PLVMQAGFLL	174	11	18	90
POL	PLPIHTAELLA	711	11	16	80
POL	PLSYQHFRKL	2	11	15	75
POL	PMGVGLSPRL	505	11	16	80
NUC	PPAYRPPNAPI	129	11	19	95
BW	POAMQMNSTTF	106	11	16	80
BW	POSLSWMTSL	192	11	18	90
X	QLDPADVQL	8	11	16	80
POL	QVFADATPTGW	685	11	19	95
POL	RLKLMPARFY	106	11	15	75
POL	RPIDWKCQRI	615	11	16	80
NUC	RPPNAPILSTL	133	11	20	100
NUC	ROALCWGELM	56	11	18	90
NUC	ROLVWRHISCL	98	11	18	90
POL	RVAEDLNIGNL	36	11	18	90
POL	RVHFFASPLHVA	818	11	15	75
POL	SIPWTHKVGNF	49	11	20	100
BW	SLDSWMTSLNF	194	11	19	95
BW	SILVPEVQWVF	337	11	19	95
NUC	SPEHCSPHHTA	44	11	20	100
NUC	SPFLAOFESA	511	11	19	95
NUC	SPHHTALROAI	49	11	20	100
BW	SPTVWLSVWMM	350	11	15	75
BW	SVRFSWLSLV	330	11	16	80
POL	SVLSRKYTSF	739	11	17	85
POL	SVRRAPPHQL	524	11	19	95
POL	TPARVTGGVFL	354	11	18	90
POL	TOCGYPALMPL	636	11	19	95
NUC	TVOASKLQGW	16	11	16	80
BW	VLDYQGMPLPV	259	11	18	90
POL	VLSRKYTSFPW	741	11	17	85
POL	VPNLSLTMLL	398	11	19	95

# HBV B62 Super Motif

Protein	Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
N/C	VOASKLQGL	17	11	16	80
BW	VQWFVGLSPTV	343	11	19	95
POL	VLGAKSVCHL	542	11	16	80
POL	VWRAFPHCLA	525	11	19	95
POL	WILGTSFVYV	759	11	16	80
POL	WILGCAAWIL	751	11	16	80
POL	WLSLDSAFY	414	11	19	95
BW	WLSLVFVGV	335	11	20	100
BW	WMCLRRHFL	237	11	19	95
BW	WMWVWVGPSTLY	359	11	17	85
POL	YLHTLWKAGIL	147	11	20	100
POL	YPLDKGKPY	122	11	20	100
POL	YPALMPLVACI	640	11	19	95

464

Table XV  
HBV A01 Motif with Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0101
100	20	POL	166	ASFCGSPY	8	
90	18	POL	737	DNSVLSRIKY	10	0.0001
95	19	POL	631	FAAPFTOCGY	10	0.0680
95	19	POL	630	GFAPFTOCGY	11	
75	15	MLC	140	GRETMEY	8	
85	17	POL	579	GYSLNFMGY	9	
100	20	POL	149	HTLWKAGILY	10	0.1100
95	19	POL	653	KOAFTEPPTY	10	0.0001
85	17	MLC	30	LLDTASALY	9	12.0000
95	19	POL	415	LSLDVSAAFY	10	0.0150
75	15	MLC	137	LTFGRETMEY	11	
85	17	EMV	360	MMWVWGPSLY	10	0.0810
75	15	X	103	MSTTLEAY	9	0.8500
90	18	POL	738	NSVLSRIKY	9	0.0005
100	20	POL	124	PLDKGKPY	9	
100	20	POL	124	PLDKGKPY	10	0.1700
85	17	POL	797	PTTGRTSLY	9	0.2100
100	20	POL	165	SASFCGSPY	9	
95	19	POL	416	SILDVSAAFY	9	5.2000
75	15	X	104	STDLEAY	8	
85	17	POL	798	TTGRTSLY	8	
95	19	POL	414	WLSLDVSAAFY	11	
85	17	EMV	359	WMWVWGPSLY	11	
95	19	POL	640	YPALMPLY	8	0.3200
85	17	POL	580	YSLNFMGY	8	



TABLE XVI  
IBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
85	17	POL	721	AACFARSR	8	
85	17	POL	721	AACFARSRSGA	11	0.0004
95	19	POL	632	AAPTOCGY	9	
95	19	POL	632	AAPTOCGYPA	11	
85	17	POL	722	ACFARSRGA	10	
80	16	POL	688	ADATPTGWGLA	11	
90	18	POL	776	ADDPGRGR	8	
95	19	POL	529	AFPHCLAF	8	
95	19	POL	529	AFPHCLAFSY	10	
95	19	POL	62	AFSSAGPCA	9	
90	18	X	62	AFSSAGPCALR	11	
95	19	POL	655	AFTFSPTY	8	
95	19	POL	655	AFTFSPTYK	9	0.2600
95	19	POL	655	AFTFSPTYKA	10	
95	19	POL	655	AFTFSPTYKAF	11	
80	16	ENW	180	AGFLTR	8	
90	18	X	66	AGPCALRF	8	
90	18	X	66	AGPCALRTSA	11	
95	19	POL	18	AGPLEELPR	10	0.0004
95	19	POL	521	AICSVRR	8	-0.0002
95	19	POL	521	AICSVVRR	9	
95	19	POL	521	AICSVVRRRA	10	
95	19	POL	521	AICSVVRRRAF	11	
95	19	NJC	41	ALESPEHCSPH	10	
90	18	POL	772	ALNPADPSR	8	0.0003
85	17	X	70	ALRFSAR	8	0.0047
80	16	ENW	108	AMOWNSTTF	9	
80	16	ENW	108	AMOWNSTTFH	10	
75	15	X	102	AMSTTDLEA	9	
85	17	NJC	34	ASALYREA	8	
100	20	POL	166	ASFCGSPY	8	0.0460
80	16	POL	822	ASPLHVAWR	9	
75	15	ENW	84	ASTINROGR	9	0.0009
80	16	POL	690	ATPTGWGLA	8	
80	16	POL	755	CAANWILR	8	
95	19	X	61	CAFSSAGPCA	10	
90	18	X	69	CALRFTSA	8	
85	17	X	69	CALRFTSAR	9	0.0034
80	16	X	6	CCQLDPAR	8	
85	17	POL	723	CFARSRSGA	9	
75	15	POL	607	CFRKL PVNR	9	
95	19	POL	638	CGYPALMPLY	10	
95	19	POL	638	CGYPALMPLYA	11	

# HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0.301
100	20	EW	312	CIPIPSSWA	9	
100	20	EW	312	CIPIPSSWAF	10	
80	16	EW	312	CIPIPSSWAFA	11	
95	19	EW	253	CLFLVLVDY	11	0.0083
90	18	X	17	CLPVGAEGR	10	0.0011
95	19	EW	239	CLRRFIF	8	
75	15	EW	239	CLRRFIFL	10	
100	20	MC	48	CSPHHTALR	9	
100	20	MC	48	CSPHHTALRQA	11	0.0029
95	19	POL	523	CSVRRRAF	8	
95	19	POL	523	CSVRRTAAPH	10	
100	20	EW	310	CTCIPIPSSWA	11	
80	16	POL	689	DAIPTGWGLA	10	
90	18	POL	540	DDVVLGAK	8	
90	18	MC	31	DIDPYKEF	8	
90	18	MC	31	DIDPYKEFGA	10	
85	17	MC	29	DLLDTASA	8	
85	17	MC	29	DLLDTASALY	10	0.0001
85	17	MC	29	DLLDTASALYR	11	0.0042
95	19	EW	196	DSWMTSLNF	9	0.0006
85	17	MC	32	DTASALYR	8	0.0004
80	16	MC	32	DTASALYREA	10	
95	19	X	14	DMCLRPVGA	10	
95	19	POL	418	DVSAFVH	8	
90	18	POL	541	DMVLGAKSVQH	11	
95	19	POL	17	EAGPLEELPR	11	
90	18	MC	40	EALSPFH	8	-0.0009
90	18	POL	718	ELLACFA	8	
90	18	POL	718	ELLACFAR	9	0.0002
85	17	POL	718	ELLACFARSR	11	0.0062
95	19	MC	43	ELSLPSPDF	10	
95	19	MC	43	ELSLPSPDF	11	
95	19	MC	43	ESPEHCSPH	9	
95	19	MC	43	ESPEHCSPHH	10	
95	19	POL	374	ESRLVDF	8	
95	19	POL	374	ESRLVDFSQF	11	
95	19	MC	174	ETTVMRRR	8	0.0003
80	16	MC	174	ETTVMRRRGR	10	0.0003
95	19	POL	631	FAAPTQCGY	10	
85	17	POL	724	FARSRSQA	8	
80	16	POL	821	FASPLHVA	8	
80	16	POL	821	FASPLHVAWR	10	

# HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
90	18	ENV	13	FFPDHQLDPA	10	
85	17	ENV	13	FFPDHQLDPAF	11	
75	15	MUC	139	FGRETVEY	9	
75	15	POL	244	FGVEPSSGSH	10	
95	19	MUC	122	FGWIRTPPA	10	
95	19	MUC	122	FGWIRTPPAY	11	
80	16	ENV	248	FILLCLIF	9	
80	16	ENV	246	FILLCLIF	11	
75	15	ENV	171	FLGPLVLQA	10	
95	19	POL	513	FLAQTSA	9	0.0006
95	19	POL	562	FLSLGIH	8	
95	19	ENV	256	FLVLIDY	8	0.0050
100	20	POL	363	FLVDKNPH	8	
95	19	POL	658	FSPTYKAF	8	
95	19	X	63	FSSAGPCA	8	
90	18	X	63	FSSAGPCALR	10	
90	18	X	63	FSSAGPCALRF	11	
100	20	ENV	333	FSYMDVVLGA	11	0.0004
90	18	POL	536	FTSPTYK	8	0.0100
95	19	POL	656	FTSPTYKA	9	
95	19	POL	656	FTSPTYKAF	10	0.0004
95	19	POL	635	FTOAGYPA	8	
95	19	POL	518	FTSACSVVR	10	0.0003
95	19	POL	518	FTSACSVVR	11	0.0065
95	19	X	132	FVLGGCRH	8	
90	18	X	132	FVLGGCRHK	9	0.0430
90	18	POL	766	FVYVPSALNPA	11	
80	16	POL	754	GCAANMILH	9	
95	19	POL	630	GFAPFTQCGY	11	
90	18	ENV	12	GFFPDHQLDPA	11	
75	15	ENV	170	GFLGPLVLQA	11	
85	17	ENV	61	GALLGWSPQA	10	
100	20	POL	360	GGVFLVDK	8	
100	20	POL	360	GGVFLVDKNPH	11	
75	15	POL	567	GHLNPNK	8	
75	15	POL	567	GHLNPNKTK	10	0.0025
75	15	POL	567	GHLNPNKTKR	11	
85	17	POL	682	GLCOVFADA	9	0.0001
95	19	POL	627	GLLGFAPF	9	0.0006
85	17	ENV	62	GILGWSPQA	9	
95	19	X	57	GLPVCAFSSA	10	

# HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
95	19	POL	509	GLSPFLA	8	
95	19	POL	509	GLSPFLAOF	10	
85	17	NUC	29	GMDIDPYK	8	0.0006
85	17	NUC	29	GMDIDPYKEF	10	-0.0003
90	18	POL	735	GTDNSVLSR	10	0.0010
90	18	POL	735	GTDNSVLSRK	11	0.0140
80	16	POL	763	GTSFVVPSPA	10	
80	16	POL	245	GVEPSGSH	9	
100	20	POL	361	GVFLVDKNPH	10	
80	16	POL	507	GVGLSPFLA	10	
95	19	NUC	123	GVWRTTPA	9	
95	19	NUC	123	GVWRTTPAY	10	0.0047
95	19	NUC	123	GVWRTTPAYR	11	0.1900
100	20	NUC	47	HCSPHHTA	8	
100	20	NUC	47	HCSPHHTALR	10	
80	16	POL	820	HFASPLHVA	9	
80	16	POL	820	HFASPLHVAWR	11	
95	19	X	49	HGAHLSLR	8	
85	17	EN	60	HGALLGWSPQA	11	
90	18	NUC	104	HISCLTGR	9	~9
75	15	POL	569	HLNPNKTK	8	
75	15	POL	569	HLNPNKTKR	9	
90	18	X	52	HLSLRLPYCA	11	
80	16	POL	491	HLYSHIPILGF	11	
85	17	POL	715	HTAELLAA	8	
85	17	POL	715	HTAELLAACF	10	
85	17	POL	715	HTAELLAA	11	
100	20	POL	149	HTLWKAGLY	10	0.0440
100	20	POL	149	HTLWKAGILYK	11	0.5400
95	19	POL	522	ICSVVRRA	8	
95	19	POL	522	ICSVVRRAAF	9	
95	19	POL	522	ICSVVRRAAPH	11	
90	18	NUC	32	IDPYKEFGA	9	
90	18	POL	617	IDMKVCOH	8	
100	20	EN	381	IFCLWVY	8	
95	19	EN	255	IFLLVLDY	9	
80	16	POL	734	IGTNSVLSR	11	
100	20	EN	249	ILLCLIF	8	
80	16	POL	760	ILBGTSPVY	9	0.0440
90	18	NUC	105	ISCLTFGR	8	0.0004
90	18	POL	625	IVGLGF	8	
90	18	POL	625	IVGLGF	9	

# HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
90	18	POL	625	IVGLGFAPF	11	
100	20	POL	153	KAGILYKR	8	0.0002
80	16	POL	503	KIPMGVGLSPF	11	
75	15	POL	108	KLIMPAPRF	8	
75	15	POL	108	KLIMPAPRF	9	
80	16	POL	610	KLPVNPAPIDWK	11	
85	17	POL	574	KTKRMWGYSLNF	11	
75	15	X	130	KVFYLGGRH	9	0.0420
75	15	X	130	KVFYLGGRH	10	
95	19	POL	55	KVGNFTGLY	9	0.2100
85	17	POL	720	LACGFARSR	9	0.0058
95	19	X	16	LCLRPVGA	8	
90	18	X	16	LCLRPVGAESH	11	
95	19	POL	683	LCQVFADA	8	
100	20	POL	125	LDKGKPY	8	
100	20	POL	125	LDKGKPY	9	
80	16	X	9	LDPARDVLCIR	11	
95	19	EW	195	LDSWWTSLNF	10	
85	17	NJC	31	LDTASALY	8	
85	17	NJC	31	LDTASALY	9	0.0004
80	16	NJC	31	LDTASALY	11	
95	19	POL	417	LDVSAFY	8	
95	19	POL	417	LDVSAFYH	9	
80	16	EW	247	LFILLCLIF	10	
95	19	POL	544	LGAKSVCH	8	
80	16	POL	753	LGCAANWILR	10	
75	15	POL	566	LGHLNPNK	9	
75	15	POL	566	LGHLNPNKTK	11	
95	19	EW	172	LGPLVLAQF	9	
95	19	EW	172	LGPLVLAQF	11	0.0022
95	19	EW	254	LIFLLVLDY	10	-0.0002
100	20	POL	109	LIMPAPFY	8	0.0024
90	18	POL	719	LLACGFAR	8	
85	17	POL	719	LLACGFARSR	10	
85	19	POL	514	LLAQFTSA	8	
95	17	NJC	30	LLDTASALY	9	0.0013
85	17	NJC	30	LLDTASALY	10	0.0050
80	16	POL	752	LLGCANWILR	11	
95	19	POL	628	LLGFAAPF	8	
85	17	EW	63	LLGWSPQA	8	
100	20	EW	378	LLPIEFCLWVY	11	
95	19	NJC	44	LLSFLPSDF	9	0.0230

# HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
95	19	NUC	44	LSFLPSDF	10	
95	19	EW	175	LLVLAQGF	8	
95	19	EW	175	LLVLAQGF	9	0.0006
100	20	EW	338	LLVPVQWF	9	
85	17	NUC	100	LLWFHISCLTF	11	
95	19	NUC	45	LSFLPSDF	8	
95	19	NUC	45	LSFLPSDF	9	0.0006
95	19	NUC	415	LSLVSAA	8	
95	19	POL	415	LSLVSAA	9	
95	19	POL	415	LSLVSAA	10	
95	19	POL	415	LSLVSAAFY	11	
95	19	POL	415	LSLVSAAFYH	11	
75	15	POL	564	LSLGHILNPNK	11	
100	20	EW	336	LSLVPVQWF	11	
95	19	X	53	LSLRLPVCA	10	
95	19	X	53	LSLRLPVCAF	11	
95	19	POL	510	LSPFLAOF	9	
95	19	POL	742	LSRKYTSF	8	
85	17	NUC	169	LSTLPETTVA	11	-0.0009
95	19	EW	16	LSVNPPLGF	9	
75	15	EW	412	LSWLSLVSAA	10	0.0048
100	20	POL	412	LSWLSLVSAA	11	
95	19	POL	3	LSYOHFRK	8	
75	15	POL	137	LTFGRETLEY	11	
75	15	NUC	99	LTVNEKR	8	-0.0002
85	17	POL	176	LVLQAGF	8	
95	19	EW	339	LVPVQWF	8	
100	20	EW	119	LVSFGWIR	9	0.0028
90	18	NUC	377	LWDFSOF	8	
100	20	POL	377	LWDFSOF	10	
100	20	POL	238	MCLRRHIF	9	0.0016
95	19	EW	238	MCLRRHIF	11	
75	15	EW	539	MDDVILGA	8	
90	18	POL	539	MDDVILGA	9	
90	18	POL	30	MDIDPYKEF	9	
90	18	NUC	30	MDIDPYKEFGA	11	
80	16	NUC	506	MGVGLSPF	8	
80	16	POL	506	MGVGLSPFLA	11	
85	17	EW	360	MMWYWGPSLY	10	0.0500
80	16	X	103	MSTIDLEA	8	
75	15	X	103	MSTIDLEAY	9	
75	15	X	103	MSTIDLEAYF	10	0.0008
75	15	X	103	MSTIDLEAYFK	11	

# HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
95	19	POL	561	NFLSLGIH	9	
90	18	NJC	75	NLEDPASR	8	-0.0002
95	19	POL	45	NLNVSIPTWH	10	
95	19	POL	45	NLNVSIPTWHK	11	-0.0009
75	15	ENV	15	NLSVNPPLGF	10	
90	18	POL	411	NLSWLSLDVSA	11	
75	15	ENV	215	NSOSPTSNH	9	0.0006
90	18	POL	738	NSVLSRK	8	0.0002
90	18	POL	738	NSVLSRK	9	
100	20	POL	47	NSVSIPTWH	8	
100	20	POL	47	NSVSIPTWHK	9	0.0820
90	18	POL	775	PADDPSSRGH	8	0.0008
95	19	POL	641	PALMPLYA	10	
75	15	X	145	PAPCNFTSA	9	0.0002
80	16	X	11	PARVLTGIVF	9	
90	18	POL	355	PASTNRQSGH	10	
75	15	ENV	83	PAYRPPNA	8	
95	19	NJC	130	PCALRTISA	9	
90	18	X	68	PCALRTISA	10	
85	17	X	68	PCALRTISA	8	
75	15	X	147	PCNFTSA	8	
95	19	ENV	15	PDHOLDPA	8	
90	18	ENV	15	PDHOLDPAF	9	
95	19	POL	512	PELUAQTS	10	
95	19	POL	634	PFTOCGYPA	9	
100	20	ENV	233	PGYRMWCLH	10	0.0008
95	19	ENV	233	PGYRMWCLFR	11	0.0048
95	19	ENV	233	PGYRMWCLFRF	9	
90	18	POL	616	PIDMKVCCOR	9	0.0002
100	20	ENV	380	PIEFLWVY	9	0.0011
85	17	POL	713	PIHTAELLA	10	
85	17	POL	713	PIHTAELLA	8	
80	16	POL	496	PILGFRK	8	
100	20	ENV	314	PIPSWAF	8	
80	16	ENV	314	PIPSWAF	9	0.0001
100	20	POL	124	PLDKGKPY	10	0.0002
100	20	POL	124	PLDKGKPY	8	0.0002
95	19	POL	20	PLEELPR	10	
90	18	POL	20	PLEELPR	8	
95	19	ENV	10	PLGFFPDH	9	
100	20	POL	427	PLHPAAMPH	9	0.0012
95	19	ENV	174	PLVLOAGF	9	

# HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
95	19	EW	174	PLVLQAGFF	10	
80	16	POL	711	PLPHITAELLA	11	
100	20	POL	2	PLSYQHRR	8	-0.0002
75	15	POL	2	PLSYOHFRK	9	0.0011
85	17	POL	98	PLTVNEKR	8	0.0002
85	17	POL	98	PLTVNEKRR	9	0.0008
80	16	POL	505	PMGVGLSPF	9	
85	17	POL	797	PTTGRTSLY	9	0.0001
85	17	POL	797	PTTGRTSLYA	10	
95	19	X	59	PVCAFSSA	8	
90	18	X	20	PVGAESRRR	9	0.0002
85	17	POL	612	PVNRPIDWK	9	0.0310
95	19	POL	654	QAFTESPTY	9	0.0030
95	19	POL	654	QAFTESPTYK	10	0.0450
80	16	EW	179	QAFTESPTYKA	11	
80	16	EW	107	QAGFELLTR	9	
80	16	EW	107	QAMQWNSTTF	10	
80	16	EW	107	QAMQWNSTTFH	11	
95	19	POL	637	OCGYPALMPY	11	
95	19	POL	517	QFTSAICSVR	11	
75	15	NJC	169	OSPRRRRSQR	11	
80	16	POL	189	QSSGLSR	8	
95	19	POL	528	RAFPICLA	8	
95	19	POL	528	RAFPICLAF	9	
95	19	POL	528	RAFPICLAFSY	11	0.0015
85	17	NJC	28	RDLLDTASA	9	0.1200
85	17	NJC	28	RDLLDTASALY	11	
95	19	X	13	RDVLCIAPVGA	11	
100	20	EW	332	RFSWLSILVPF	11	
95	19	X	56	RGLPVCAF	8	
95	19	X	56	RGLPVCAFSSA	11	
100	20	NJC	152	RGRSPRRR	8	
80	16	POL	762	RGTSFVYVPSA	11	
90	18	POL	624	RIVGLIGF	8	
90	18	POL	624	RIVGLIGFA	9	
90	18	POL	624	RIVGLIGFAA	10	
75	15	POL	106	RLKLMPA	8	
75	15	POL	106	RLKLMPAR	9	0.0950
75	15	POL	106	RLKLMPARF	10	
75	15	POL	106	RLKLMPARFY	11	
75	15	X	128	RLKVFVLGGR	11	
95	19	POL	376	RLVDESCF	9	0.0006



# HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
95	19	POL	376	RLVDFSCFSR	11	0.2800
95	19	MJC	183	RSPRRRTSPR	11	-0.0007
75	15	MJC	167	RSQSPRRR	8	
75	15	MJC	167	RSQSPRRR	9	
90	18	POL	353	RTPARVTGVF	11	
95	19	MJC	127	RTPPAYRPNNA	11	
95	19	MJC	188	RTSPRRRR	8	-0.0002
95	19	MJC	188	RTSPRRRR	9	0.0054
80	16	POL	818	RVHFAAPLH	9	
75	15	POL	818	RVHFAAPLHVA	11	
100	20	POL	357	RVTGVEFLVDK	11	0.0190
90	18	X	65	SAGPCALR	8	-0.0002
90	18	X	65	SAGPCALRF	9	-0.0003
95	19	POL	520	SAICSVVR	8	-0.0002
95	19	POL	520	SAICSVVR	9	0.0058
95	19	POL	520	SAICSVVRR	10	
95	19	POL	520	SAICSVVRR	11	
95	19	POL	520	SAICSVVRR	11	
90	18	POL	771	SAINPADDPSSR	9	-0.0004
100	20	POL	165	SASFCGSPY	11	
90	18	MJC	121	SFGWIRITPPA	11	
95	19	MJC	46	SFLPSDF	8	
75	15	POL	748	SFPWLLGCA	9	
75	15	POL	748	SFPWLLGCA	10	
80	16	POL	765	SFVYVPSA	8	
100	20	POL	49	SIPWTHKGVNF	11	
95	19	EMV	194	SIDSMWTSLSNF	11	
95	19	POL	416	SIDVSAAF	8	
95	19	POL	416	SIDVSAAFY	9	0.0016
95	19	POL	416	SIDVSAAFYH	10	
75	15	POL	565	SLGHLNPNK	10	
100	20	EMV	337	SLVPRVQWF	10	
95	19	X	54	SLRGLPVCA	9	
95	19	X	54	SLRGLPVCAF	10	0.0004
90	18	X	64	SSAGPCALR	9	0.0080
90	18	X	64	SSAGPCALRF	10	-0.0003
95	19	MJC	170	STLPETTIVRR	10	0.0007
80	16	EMV	85	STNROSGR	11	0.0150
75	15	X	104	STDLEAV	8	
75	15	X	104	STDLEAYF	9	
75	15	X	104	STDLEAYFK	10	
75	15	EMV	17	SNPRLGF	8	0.0066

### HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
90	18	POL	739	SWLSRKYY	8	-0.0002
85	17	POL	739	SWLSRKYYTSF	11	
95	19	POL	524	SVVRRAPFH	9	0.1100
85	17	POL	716	TAELLAACFA	9	
85	17	POL	716	TAELLAACFA	10	
85	17	POL	716	TAELLAACFAR	11	0.0006
80	16	MJC	33	TASALYREA	9	
100	20	ENW	311	TCIPIPSWA	10	
100	20	ENW	311	TCIPIPSWAF	11	
80	16	X	106	TDLEAYFK	8	
90	18	POL	736	TDNSVLSR	9	
90	18	POL	736	TDNSVLSRK	10	0.0006
90	18	POL	736	TDNSVLSRKYY	11	
75	15	MJC	138	TFGRETVLEY	10	
95	19	POL	657	TFSPITYKA	8	
95	19	POL	657	TFSPITYKAF	9	
100	20	POL	359	TGQWELVDK	9	0.0007
85	17	POL	799	TGRTSLYA	8	
95	19	MJC	171	TLPETTVR	9	0.0008
95	19	MJC	171	TLPETTVRR	10	0.0007
95	19	MJC	171	TLPETTVRRR	11	0.0005
100	20	POL	150	TLWKAGILYK	9	0.1300
100	20	POL	150	TLWKAGILY	10	5.3000
100	20	POL	150	TLWKAGILYK	11	0.0082
95	19	POL	519	TSAICSVVR	9	0.0005
95	19	POL	519	TSAICSVRR	10	0.0018
95	19	POL	519	TSAICSVVRR	11	
75	15	POL	747	TSFPMILGCA	10	
75	15	POL	747	TSFPMILGCAA	11	
80	16	POL	764	TSFVVPSPA	9	
75	15	X	105	TTDLEAYF	8	
75	15	X	105	TTDLEAYFK	9	0.0006
85	17	POL	798	TTGRTSLY	8	0.0004
85	17	POL	798	TTGRTSLYA	9	
75	15	ENW	278	TTSTGPCK	8	
80	16	MJC	175	TTVRRRGR	9	0.0008
80	16	MJC	176	TTVRRRGR	8	0.0003
80	16	MJC	176	TVRRRGR	11	
95	17	X	60	TVRRRGRSPR	11	
85	19	POL	621	VCAFSAGPCA	11	
100	20	POL	379	VOCIRVGLGF	11	
100	20	POL	379	VDFSQFSR	8	
100	20	POL	362	VFLVDKNPH	9	

# HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
80	16	X	131	VFALGGCR	8	
80	16	X	131	VFALGGCRH	9	
75	15	X	131	VFALGGCRHK	10	
95	19	X	21	VGAESRGR	8	
95	19	POL	626	VGILGFAA	8	
95	19	POL	626	VGLLGFAPF	10	
80	16	POL	508	VGLSPFLA	9	
80	16	POL	508	VGLSPFLAQF	11	
95	19	POL	56	VGNFTGLY	8	
85	17	POL	96	VGPLTVNEK	9	
85	17	POL	96	VGPLTVNEKR	10	0.0007
85	17	POL	96	VGPLTVNEKRR	11	
95	19	X	15	VLCLPYGA	9	
95	19	POL	543	VLGAKSVQH	9	
90	18	X	133	VLGGCRHK	8	0.0150
80	16	ENV	177	VLQAGFLLTR	11	
85	17	POL	741	VLNRKYTSF	9	
90	18	MJC	120	VSEFGWIR	8	0.0040
100	20	POL	48	VSIPTWTHK	8	0.0130
100	20	POL	358	VTGAVFLVDK	10	0.0390
100	20	POL	378	WDFSQFSR	9	0.0015
90	18	POL	542	VLGAKSVQH	10	
85	17	POL	740	VLNRKYTSF	10	0.0004
95	19	POL	525	VVRRAPHL	8	
95	19	POL	525	VVRRAPHLA	11	
80	16	MJC	177	VVRRGRSPRR	10	0.0027
80	16	MJC	177	VVRRGRSPRR	11	
90	18	MJC	102	WFHISCLTF	9	
90	18	MJC	102	WFHISCLTFGR	11	
85	17	MJC	28	WGMDIDPY	8	
85	17	MJC	28	WGMDIDPYK	9	-0.0003
85	17	MJC	28	WGMDIDPYKEF	11	
85	17	POL	578	WGSUNFMGY	10	
80	16	POL	759	WLIGTSF	8	
80	16	POL	759	WLIGTSFVY	10	0.0076
95	19	MJC	125	WIRTPPAY	8	-0.0002
95	19	MJC	125	WIRTPPAYR	9	0.0008
95	18	POL	314	WLOFRNSK	8	
100	20	POL	414	WLSLDVSA	8	-0.0002
95	19	POL	414	WLSLDVSAA	9	
95	19	POL	414	WLSLDVSAAF	10	
95	19	POL	414	WLSLDVSAAFY	11	0.0034

# HBV A03 Motif With Binding Data

Conservancy	Freq.	Protein	Position	Sequence	AA	A*0301
100	20	ENW	335	WLSLLVPE	8	
85	17	NJC	26	WLWGMDDIPY	10	0.0002
85	17	NJC	26	WLWGMDDIPYK	11	0.0030
95	19	ENW	237	WMCLRRFIIF	10	0.0004
85	17	ENW	359	WMWMMWVGPSLY	11	0.0009
100	20	POL	52	WTHKXGNF	8	
100	20	POL	147	YLHTLWKA	8	
100	20	POL	122	YLPIDKGIK	9	0.0001
100	20	POL	122	YLPIDKGIKPY	11	-0.0004
90	18	NJC	118	YLVSEGWIR	10	0.0005
90	18	POL	538	YMDVVLGA	9	0.0001
90	18	POL	538	YMDVVLGAK	10	0.0330
80	16	POL	493	YSHPIILGF	9	
80	16	POL	493	YSHPIILGFR	10	
80	16	POL	493	YSHPIILGFRK	11	
85	17	POL	580	YSLNFMGY	8	
75	15	POL	746	YTSFPWLLGCA	11	
90	18	POL	768	YVPSALNPA	9	
				480		

-0.0002

Table XVII  
 All Motif With Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*1101
85	17	POL	721	AACFARSR	8	
95	19	POL	632	AAPTOCGY	9	
90	18	POL	776	ADDPGRGR	8	
95	19	POL	529	AFPHCLAFSY	110	
90	18	X	62	AFSSAGPCALR	11	
95	19	POL	655	AFTFSPTY	8	
95	19	POL	655	AFTFSPTYK	9	
80	16	ENW	180	AGFLTLTR	8	
95	19	POL	18	AGPLEEELPR	10	
95	19	POL	521	AICSVRR	8	
95	19	NJC	41	ALSPHICSPH	11	
90	18	POL	772	ALNPADPSR	10	
85	17	X	70	ALRFTSR	8	
80	16	ENW	108	AMQWNSITFH	10	
100	20	POL	166	ASFCSPY	8	
80	16	POL	822	ASPLHAWRH	9	
75	15	ENW	84	ASTNRQSGR	9	
80	16	POL	755	CAANNILR	8	
85	17	X	69	CALRFTSAR	9	
80	16	X	6	CCOLDPAR	8	
75	15	POL	607	CFRKLPMNR	9	
95	19	POL	638	CGYPALMPLY	10	
95	19	ENW	253	CLIFELVILDY	11	
90	18	X	17	CLPVGAESE	10	
100	20	NJC	48	CSPHHTALR	9	
95	19	POL	523	CSVRRRAFPH	10	
90	18	POL	540	DDVLGAK	8	
85	17	NJC	29	DLLDTASALY	10	
85	17	NJC	29	DLLDTASALY	11	
90	18	POL	737	DNSVLSR	8	
90	18	POL	737	DNSVLSRK	9	
90	18	POL	737	DNSVLSRKY	10	
85	17	NJC	32	DTASALYR	8	
95	19	POL	418	DVSAAPYH	8	
90	18	POL	541	DVLGAKSVOH	11	
95	19	POL	17	EAGPLEELPR	11	
90	18	NJC	40	EALSPRH	8	
90	18	POL	718	ELLAACFAR	9	
85	17	POL	718	ELLAACFARSR	11	
95	19	NJC	43	ESPEHCSPH	9	
95	19	NJC	43	ESPEHCSPH	10	
95	19	NJC	174	ETTIVRRR	8	

# A11 Motif With Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*1101
80	16	MJC	174	ETVRRRRGR	10	
95	19	POL	631	FAAPFTOCGY	10	
80	16	POL	821	FASPLHVAWR	10	
75	15	MJC	139	FGRETVLEY	10	
75	15	POL	244	FGVEPSGSGH	9	
95	19	MJC	122	FGWIRTPPAY	10	
95	19	POL	562	FLSLGIH	11	
95	19	BNW	256	FLVLVDY	8	
100	20	POL	363	FLVDKNPH	8	
90	18	X	63	FSSAGPCALR	8	
95	19	POL	656	FTFSPTYK	10	
95	19	POL	518	FTSAICSVR	8	
95	19	POL	518	FTSAICSVRR	10	
95	19	X	132	FVLGGCH	11	
90	18	X	132	FVLGGCHK	8	
80	16	POL	754	GCAANMLR	9	
95	19	POL	630	GEAPFTOCGY	11	
100	20	POL	360	GGVFLVDK	8	
100	20	POL	360	GGVFLVDKNPH	11	
75	15	POL	567	GHLNPNK	8	
75	15	POL	567	GHLNPNKTK	10	
75	15	POL	567	GHLNPNKTKR	11	
85	17	MJC	29	GMDIDPYK	8	
95	19	POL	44	GNLNSIPWTH	11	
90	18	POL	735	GTDNSWLSR	10	
90	18	POL	735	GTDNSWLSRK	11	
80	16	POL	245	GVEPSGSGH	9	
100	20	POL	361	GVFLVDKNPH	10	
95	19	MJC	123	GWIRTPPAY	10	
95	19	MJC	123	GWIRTPPAYR	11	
100	20	MJC	47	HCSPHHTALR	10	
80	16	POL	820	HFAAPLHVAVR	11	
95	19	X	49	HGAHLSLR	8	
90	18	MJC	104	HISCLTEGR	8	
75	15	POL	569	HLNPNKTK	9	
75	15	POL	569	HLNPNKTKR	8	
100	20	POL	149	HTLWKAGILY	9	
100	20	POL	149	HTLWKAGILYK	10	
95	19	POL	522	ICSVVRRAEPH	11	
90	18	POL	617	IDWKVQCR	11	
100	20	BNW	381	IFCLWVY	8	
95	19	BNW	255	IFLVLDY	9	

# All Motif With Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*1101
80	16	POL	734	IGDINSVLSR	11	
80	16	POL	760	ILRGTSFVY	9	
90	18	MJC	105	ISCLTFGR	8	
100	20	POL	153	KAGILYKR	8	
75	15	POL	108	KLIMPARFY	9	
80	16	POL	610	KLPVNRPIDWK	11	
75	15	X	130	KFVILGGCR	9	
75	15	X	130	KFVILGGCRH	10	
95	19	POL	55	KVGNFTGLY	9	
85	17	POL	720	LAACFATSR	9	
90	18	X	16	LCLIRPYGAESH	11	
100	20	POL	125	LDKIKIPY	8	
100	20	POL	125	LDKIKIPY	9	
80	16	X	9	LDPARDYLCLR	11	
85	17	MJC	31	LDTSALY	8	
85	17	MJC	31	LDTSALYR	9	
95	19	POL	417	LDVSAFVY	8	
95	19	POL	417	LDVSAFVYH	9	
95	19	POL	544	LGAKSVQH	8	
80	16	POL	753	LGCAANWILR	10	
75	15	POL	566	LGIHLNPNK	9	
75	15	POL	566	LGIHLNPNKTK	11	
95	19	ENV	254	LHLLVLDY	10	
100	20	POL	109	LIMPARFY	8	
90	18	POL	719	LLACFAR	8	
85	17	POL	719	LLACFARSR	10	
85	17	MJC	30	LLDTASALY	9	
85	17	MJC	30	LLDTASALYR	10	
80	16	POL	752	LLGCANWILR	11	
100	20	ENV	378	LLPIFFCLWVY	11	
90	18	POL	773	LNPADDPGR	11	
90	18	POL	773	LNPADDPGRH	9	
75	15	POL	570	LNPKNKTKR	8	
75	15	POL	570	LNPKNKTKRWGY	11	
95	19	POL	46	LNVSPWTH	9	
95	19	POL	46	LNVSPWTHK	10	
95	19	POL	415	LSLDVSAFVY	10	
95	19	POL	415	LSLDVSAFVYH	11	
75	15	POL	564	LSLGIHLNPNK	11	
95	19	MJC	169	LSLPEITVVR	11	
75	15	POL	3	LSYQHETIK	8	
75	15	MJC	137	LTFGRETVLEY	11	

# A11 Motif With Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*1101
85	17	POL	99	LTVNEKRR	8	
90	18	MJC	119	LVSFGWIR	9	
100	20	POL	377	LVDVSOFSR	10	
90	18	POL	539	MDDVILGAK	9	
85	17	EW	360	MMWWGSPSLY	10	
75	15	X	103	MSTIDLEAY	9	
75	15	X	103	MSTIDLEAYFK	11	
95	19	POL	561	NFLSLGIH	9	
90	18	MJC	75	NLEDPASR	8	
95	19	POL	45	NLNVSPWTH	10	
95	19	POL	45	NLNVSPWTHK	11	
75	15	EW	215	NSQSPTSNH	9	
90	18	POL	738	NSVLSRIK	8	
90	18	POL	738	NSVLSRIKY	9	
100	20	POL	47	NVSIPTWH	8	
100	20	POL	47	NVSIPTWHK	9	
80	18	POL	775	PADDPGRGR	9	
75	16	X	11	PARDVCLR	9	
85	15	EW	83	PASTNHOSGR	10	
100	17	X	68	PCALFTSAR	10	
100	20	EW	233	PGYHWMCLR	9	
95	19	EW	233	PGYHWMCLRR	10	
90	18	POL	616	PIDWKCQR	9	
100	20	EW	380	PIEQLWVY	9	
80	16	POL	496	PILGFRK	8	
100	20	POL	124	PLDKGKPY	9	
100	20	POL	124	PLDKGKPY	10	
95	19	POL	20	PLEELPR	8	
95	19	EW	10	PLGFPDH	8	
100	20	POL	427	PLHPAAMPH	9	
100	20	POL	2	PLSYOHR	8	
75	15	POL	2	PLSYOHRK	9	
85	17	POL	98	PLTVNEKR	8	
85	17	POL	98	PLTVNEKRR	9	
75	15	POL	572	PNKTRWGY	9	
85	17	POL	797	PTTGITSLY	9	
90	18	X	20	PVGAESRGR	9	
85	17	POL	612	PVNRPIDWK	9	
95	19	POL	654	QAFTEPTV	9	
95	19	POL	654	QAFTEPTYK	10	
80	16	EW	179	QAGFLTR	9	
80	16	EW	107	QAMQWNSTFH	11	



# A11 Motif With Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*1101
95	19	POL	637	OCGYPALMPLY	11	
95	19	POL	517	OFTSAICSVVR	11	
75	15	MJC	169	OSPRRRRSQSR	11	
80	16	POL	189	OSSGLSR	11	
95	19	POL	528	FAFPHCLAFSY	18	
85	17	MJC	28	RDLDITASALY	11	
100	20	MJC	152	FGSRPRRR	11	
75	15	POL	106	RLKLIMPAR	8	
75	15	POL	106	RLKLIMPARFY	9	
75	15	POL	128	RLKVFALGGCR	11	
75	15	X	376	RLVDFSCFSR	11	
95	19	POL	183	RSPIRRITPSPH	11	
95	19	MJC	167	RSQSPRRR	11	3
75	15	MJC	167	RSQSPRRRR	8	
95	19	MJC	188	RTPSPRRR	9	
95	19	MJC	188	RTPSPRRRR	8	
80	16	POL	818	RMHFASPLH	9	
100	20	POL	357	RVTGVFLVDK	9	
90	18	X	65	SAGPCALR	11	
95	19	POL	520	SAICSVVR	8	
95	19	POL	520	SAICSVRR	8	
90	18	POL	771	SALNPADDP SR	9	
100	20	POL	165	SASFQSPY	11	
95	19	POL	416	SLDVSAEFY	9	
95	19	POL	416	SLDVSAAFYH	10	
75	15	POL	565	SLGHLNPNK	10	
90	18	X	64	SSAGPCALR	9	
95	19	MJC	170	STLPETTVR	10	
95	19	MJC	170	STLPETTVRR	11	
80	16	EW	85	STNRQGR	8	1
75	15	X	104	STIDLEAY	8	
75	15	X	104	STIDLEAYFK	8	
90	18	POL	739	SVLISRY	10	2
95	19	POL	524	SVRRAPFH	8	
85	17	POL	716	TAELLACFAR	9	
80	16	X	106	TDLEAYEK	11	
90	18	POL	736	TDNSVLSR	8	
90	18	POL	736	TDNSVLSRK	9	
90	18	POL	736	TDNSVLSRKY	10	
75	15	MJC	138	TFGRETVLEY	11	
100	20	POL	359	TGGVFLVDK	10	
95	19	MJC	171	TLPETTVR	9	

# All Motif With Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*1101
95	19	MJC	171	TLPETTVRR	10	
95	19	MJC	171	TLPETTVRRR	11	
100	20	POL	150	TLWKAGILY	9	
100	20	POL	150	TLWKAGILYK	10	
100	20	POL	150	TLWKAGILYKR	11	
95	19	POL	560	TNELLISLGIH	10	
95	19	POL	519	TSAICSVVR	9	
95	19	POL	519	TSAICSVRR	10	
75	15	X	105	TTDLAYFK	9	
85	17	POL	798	TTGRTSLY	8	
75	15	ENV	278	TTSTGPCK	8	
80	16	MJC	175	TTWTRRRGR	9	
80	16	MJC	176	TVARRRGR	8	
80	16	MJC	176	TVRRRRGRSPR	11	
100	20	POL	379	VDFOFSR	8	
100	20	POL	362	VELVDKNPH	9	
80	16	X	131	VFLGGCR	8	
80	16	X	131	VFLGGCRH	9	
75	15	X	131	VFLGGCRHK	10	
95	19	X	21	VGAESRRGR	8	
95	19	POL	56	VGNFTGLY	8	
85	17	POL	96	VGPLTVNEK	9	
85	17	POL	96	VGPLTVNEKR	10	
95	19	POL	96	VGPLTVNEKRR	11	
90	18	X	543	VLGAKSVCH	9	
80	16	ENV	133	VLGGRHK	8	
85	17	POL	177	VLOAGFFLLTR	11	
90	18	MJC	613	VNRPIDWK	8	
100	20	POL	120	VSGWMIR	8	
100	20	POL	48	VSLPWTHK	8	
100	20	POL	358	VTGGVFLVDK	10	
90	18	POL	378	VDFSOFSR	9	
90	18	POL	542	VVLGAKSVCH	10	
95	19	POL	525	VVRTRAFPH	8	
80	16	MJC	177	VRRRRGRSPR	10	
80	16	MJC	177	VRRRRGRSPRR	11	
90	18	MJC	102	WPHISCLTGR	11	
85	17	MJC	28	WGMDIDPY	8	
85	17	MJC	28	WGMDIDPYK	9	
80	16	POL	578	WGYSLNFMGY	10	
80	16	POL	759	WILRGTSFY	10	
95	19	MJC	125	WIRTPPAY	8	

# All Motif With Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*1101
95	19	NUC	125	WIRTPPAYR	9	
90	18	POL	314	WLQFFNSK	8	
95	19	POL	414	WLSLDVSAFY	11	
85	17	NUC	26	WLWGMDIDPY	10	
85	17	NUC	26	WLWGMDIDPYK	11	
85	17	EW	359	WMMWYWGPSLY	11	
100	20	POL	122	YLPIDKGIK	9	
100	20	POL	122	YLPIDKGIKPY	11	
90	18	NUC	118	YLVSGVWIR	10	
90	18	POL	538	YMDVWLGAk	10	
80	16	POL	493	YSHIPILGFR	10	
80	16	POL	493	YSHIPILGFRK	11	
85	17	POL	580	YSLNFMGY	8	
				265		

[illegible]

Conservancy	Freq.	Protein	Position	Sequence	AA	A.2401
95	19	POL	529	AFPHCLAF	8	
95	19	X	62	AFSSAGPCAL	10	0.0012
90	18	POL	535	AFSYMDDVVL	10	0.0009
95	19	POL	655	AFTFSPITYKAF	11	
80	16	ENV	108	AMQWNSTTF	9	
100	20	MUC	131	AYRPPNAPI	9	0.0310
100	20	MUC	131	AYRPPNAPIL	10	0.0042
75	15	POL	607	CFRKLVPVNRPI	11	
85	17	POL	618	DMKVCORRI	8	
85	17	POL	618	DMKVCORIVGL	11	
90	18	ENV	262	DYOGMLPVCPPL	11	0.0002
90	18	MUC	117	EYLVSGWV	9	
90	18	MUC	117	EYLVSGWVI	10	
100	20	ENV	382	FFCLWVVI	8	
80	16	ENV	182	FFLLTRIL	8	
80	16	ENV	182	FFLLTRILTI	10	
85	17	ENV	13	FFPDHOLDPAF	11	
80	16	ENV	181	GFELLTRIL	8	
80	16	ENV	181	GFELLTRILI	9	
80	16	ENV	181	GFELLTRILTI	11	
95	19	ENV	12	GFPDHQL	8	
75	15	ENV	170	GLGLPLVL	9	
80	16	POL	500	GFRKIPMGVGL	11	
85	17	MUC	29	GMDIDPYKEF	10	
90	18	ENV	265	GMLPVCPPL	8	
85	17	MUC	25	GMLWGMDI	8	
85	17	ENV	65	GWSPQAOGL	9	0.0024
85	17	ENV	65	GWSPQAOGL	10	0.0003
95	19	POL	639	GYPALMPL	8	
95	19	ENV	234	GYRWMCILRRF	10	0.0007
95	19	ENV	234	GYRWMCILRRFI	11	
75	15	POL	579	GYSILNFMGYVI	11	
80	16	POL	820	HFASPLHVAV	10	
75	15	POL	7	HERKILL	8	
100	20	POL	146	HYLHTLWKAGI	11	
100	20	ENV	381	IFECQWVVI	9	0.0087
80	16	ENV	245	IFELL	8	
80	16	ENV	245	IFELLCL	10	
80	16	ENV	245	IFELLCLLI	11	
85	17	ENV	358	IMMMWYWGPSL	11	0.0004
95	19	POL	395	KFAPNLOSL	10	0.0020
100	20	POL	121	KYLPLDKGI	9	

HBV A24 Motif With Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*2401
85	17	POL	745	KTSPFWL	8	
85	17	POL	745	KYTSPFWL	9	5.3000
80	16	ENV	247	LFILLCL	8	
80	16	ENV	247	LFILLCLI	9	
80	16	ENV	247	LFILLCLIF	10	
80	16	ENV	247	LFILLCLIFL	11	
95	19	POL	643	LMPLYACI	8	
90	18	MUC	101	LWFHISCL	8	
85	17	MUC	101	LWFHISCLTF	10	
80	16	POL	492	LYSHPIIL	8	
80	16	POL	492	LYSHPIILGF	10	1.1000
85	17	ENV	360	MMWVWGPSL	9	0.0060
85	17	ENV	361	MMWVWGPSL	8	0.0005
95	19	POL	561	NFLSLGI	8	
95	19	POL	561	NFLSLGHL	10	0.0099
80	16	POL	758	NMLRGTSF	9	
95	19	POL	512	PELLAQFTSAI	11	
95	19	POL	634	PFTOCGPAL	10	0.0002
95	19	ENV	341	PRQWFWGL	9	0.0003
80	16	POL	505	PMGVGLSPF	9	
80	16	POL	505	PMGVGLSPFL	10	
80	16	POL	505	PMGVGLSPFL	11	
80	16	POL	750	PWLLGCAANW	10	
80	16	POL	750	PWLLGCAANWI	11	
100	20	POL	51	PWTHKVGNF	9	0.0290
95	19	ENV	344	QWFEVGLSPTW	11	
75	15	ENV	242	RFIFLFI	8	
75	15	ENV	242	RFIFLFI	9	
75	15	ENV	242	RFIFLFI	10	
75	15	ENV	242	RFIFLFI	11	
100	20	ENV	332	RFESWLSLL	8	
100	20	ENV	332	RFESWLSLVPF	11	
85	17	POL	577	RMGYSLNF	8	
95	19	ENV	236	RMWCLRRF	8	0.0710
95	19	ENV	236	RMWCLRRFI	9	1.1000
95	19	ENV	236	RMWCLRRFII	10	
95	19	ENV	236	RMWCLRRFIIF	11	
100	20	POL	167	SFOGSPYSW	9	0.0710
95	19	MUC	46	SFLPSDF	8	
80	16	POL	765	SFVYVPSAL	9	
95	19	POL	413	SWLSLVSAAF	11	
100	20	ENV	334	SWLSLVPF	9	0.3900

# HBV A24 Motif With Binding Information

Conservancy	Freq.	Protein	Position	Sequence	AA	A*2401
95	19	POL	392	SWPKFAVPNL	10	5.6000
100	20	EW	197	SWWTSLNLF	8	
95	19	EW	197	SWWTSLNFL	9	0.3800
90	18	POL	537	SYMDVVL	8	
75	15	POL	4	SYOHFRKL	8	
75	15	POL	4	SYOHFRKLL	9	0.0051
75	15	POL	4	SYOHFRKLL	10	0.0660
75	15	POL	4	SYOHFRKLL	11	
75	15	MLC	138	TFGRETVL	8	
75	15	MLC	138	TFGRETVLEVL	11	
95	19	POL	657	TFSPITYKAF	9	0.0060
95	19	POL	657	TFSPITYKAF	10	0.0043
95	19	POL	686	VFADATPTGW	10	0.0180
75	15	X	131	VFALGGCHKL	11	
90	18	MLC	102	WFHISCLTF	9	0.0300
95	19	EW	345	WFGLSPTVW	10	0.0120
95	19	EW	345	WFGLSPTVWL	11	
95	19	EW	237	WMCLRRFI	8	
95	19	EW	237	WMCLRRFII	9	
95	19	EW	237	WMCLRRFIIF	10	0.0013
95	19	EW	237	WMCLRRFIIF	11	
85	17	EW	359	WMMMWVWGPSL	10	
95	19	EW	198	WWTSLNFL	8	
				108	3	

Table XIXa  
HBV DR-Super Motif

Protein	Core Sequence	Core Freq	Core Conservancy (%)	Exemplary Sequence	Position In HBV Poly- Protein	Exemplary Sequence Frequency	Exemplary Sequence Conservancy (%)
POL	FAAFITQGG	19	95	LLGFAPFTQGGYPA	628	19	95
POL	FADATPTGW	19	95	COVFADATPTGWGLA	684	16	80
POL	FAPNLOS	19	95	WPKFAPNLOSITNL	393	19	95
NJC	FGHEMLEY	15	75	CLTGHETLELYVS	136	14	70
POL	FGVETGGG	15	75	FRSGVETGGGCHD	252	6	30
NJC	FHSQITFG	18	90	LLMFHSQITFGHET	100	17	85
NJC	FHLQUSC	16	80	MOFLHQLUSCSP	1	10	50
ENV	FILLQIF	16	80	IFILLQIFLLV	245	16	80
ENV	FLLQIF	16	80	FIIFILLQIFLL	243	16	80
ENV	FLGRLVVO	15	75	TSFRLGRLVVOXGF	168	15	75
ENV	FLITRLTI	16	80	AGFLLITRLTIPOS	180	16	80
ENV	FLVLLDYO	19	95	CLIFLLVLLDYQML	253	19	95
ENV	FPVQSSSG	15	75	CLVTFPQSSSGSTVN	127	11	55
ENV	FPDLOLPA	18	90	LGHTFDLOLDPNGA	22	9	45
POL	FPYKLFVSY	19	95	FRVTFYKLFVSYMCD	527	19	95
POL	FTKPMGVG	16	80	ILGTRKPMGVGSP	498	13	65
POL	FTKQVFNTP	16	80	KQCFRKLQVFNTPDW	616	9	45
X	FSSGFCAL	19	95	VCAFSSGFCALFET	60	18	90
ENV	FSWLILVLP	20	100	SVERFSWLILVPEVO	330	16	80
POL	FTTSPTVKA	19	95	KQAFITTSPTVKAELC	653	12	60
POL	FTGLYSTIV	18	90	VGNFTGLYSTIVPF	56	11	55
ENV	FTSAICSIV	19	95	LAOFITSAICSVMRA	515	19	95
ENV	PVGLSPTVW	19	95	VQMPVGLSPTVWMLSV	343	14	70
X	PALGGCRHK	18	90	LKMPVGLGGCRKLWC	129	14	70
ENV	PVQVNVALS	19	95	LVPFQVNVFVGLSPD	339	19	95
POL	PVYVPSAN	18	90	GTSFVYVPSANPAD	763	16	80
POL	IDMVCORL	17	85	NFTIDMVCORIVGL	614	16	80
ENV	IFLILLC	16	80	FRIFLILLCJIF	242	15	75
ENV	IFLVLLDY	19	95	LCUFLVLVDYQGM	252	19	95
POL	IGTDSNWL	16	80	AKUGTDSNWSHKK	731	13	65
POL	HTAELVLA	17	85	PLPHHTAELVLCFA	711	16	80
ENV	HFILL	16	80	FRHFILLICU	241	15	75
ENV	ILLQIFL	20	100	FLILLQIFLLVLDY	246	16	80
POL	ILRGTFVY	16	80	AMMLILRGTFVYPS	757	16	80
NJC	ILSTLPETT	20	100	NAPILSTLPETTVMR	165	19	95
ENV	IPFSSWAF	20	100	CTCIPFSSWAFRF	321	8	40
NJC	ITPTVATP	19	85	GWMITPTVATPAPNA	123	19	95
POL	LMCAFRTSH	17	85	AELLMCAFRTSRSGA	717	16	80
POL	LAFTSYMDV	18	90	PHCLAFSYMDVDVIG	531	18	90
POL	LAOFTSAC	19	95	PFLAOFITSACSVV	512	19	95
NJC	LQIGMLWGM	17	85	ASKLQIGMLWGMDD	19	17	85
ENV	LQIFLLV	20	100	ILLQIFLLVLLDY	249	19	95
X	LQLPVGAE	19	95	FDVLQLPVGAESTG	13	18	90
POL	LCOVFADAT	19	95	FRGLCOVFADATPTG	680	11	55
ENV	LDSWMTSLN	19	95	POS.LDSWMTSLNFIQ	192	17	85
NJC	LDTSALYH	17	85	RDLLDTASALYREAL	28	16	80
POL	LDVSAFVH	19	95	WLSLDVSAFVHPL	425	11	55
ENV	LDYQGMALPV	18	90	LVLDYQGMALPVCFI	258	18	90
POL	LEELPILA	18	90	AGPLEELPILADEG	18	13	65
ENV	UFLILCU	16	80	IFLILILCUFL	244	16	80
POL	LGAKSVQHL	17	85	DVNLGAKSVQHLESL	541	16	80
POL	LGFAFTO	19	95	VGLGFAFTQGGY	626	19	95
POL	LGTRFMAG	19	95	PILGTRFMAGVQL	496	13	65
POL	LGNLNLSIP	19	95	DNLGNLNSIPWTH	40	19	95

## HBV DR-Super Motif:

Protein	Core Sequence	Core Freq.	Core Conservancy (%)	Exemplary Sequence	Position In HBV Poly-Protein	Exemplary Sequence Frequency	Exemplary Sequence Conservancy (%)
ENV	IGPIVIAOA	19	95	SGRIPIVIAOGFF	169	15	75
POL	LHPAMPHL	20	100	HLPHPAMPHLVG	425	9	45
ENV	LIPLVLID	19	95	LLQLIPLVLIDYOG	251	19	19
POL	UKLMPAF	15	75	KPKLUMPAFEPVN	104	7	35
X	LVNPLGGC	15	75	EIRLVPLGGCRHK	126	13	65
POL	LLAOFISAI	19	95	SPPLAOFISACSV	511	19	95
NLC	LLDTASALY	17	85	IRDLDTASALYREA	56	9	45
POL	LLGCANMI	16	80	FPMLIGCANMILBQ	749	15	75
POL	LLGFAPPT	19	95	IVGLGFAPPTDGG	625	18	90
ENV	LLGWSPOAO	17	85	HGALLGWSPOAOGL	60	15	75
ENV	LLUCIFLL	20	100	SVELLSTLPDSPTFS	247	16	80
NLC	LLSTLPSCDF	19	95	TNILLSTLGRILNPNK	41	11	55
POL	LLSLGILIN	19	95	LTNILLSSNL.SWL.SLD	560	15	75
ENV	LLTRILTP	16	80	GFLLITLITLTIPOSL	404	16	80
ENV	LLVLONGGF	19	95	LGPHLLVLONGFHLLT	181	172	90
ENV	LLVFEVOMF	20	100	WLSILVFEVOMFVGL	335	19	95
NLC	LLVPHISCL	18	90	IRQLLVPHISCLITGG	126	13	65
POL	LVAPLYACIO	19	95	YVALMPLVYACIOSQ	640	11	55
POL	LVNLGNLNV	19	95	AEDILNLGNLNVSPW	38	19	95
POL	LVNPKTKRW	15	75	GHLNPKTKRWGYS	567	15	75
POL	LVNRPVATH	17	85	DEGLNRPVATHDNLG	30	12	60
NLC	LVPTTVVRA	19	95	LGNLVSPVTHKVG	43	19	95
ENV	LVPIFCILW	20	100	LSLTPETTVVRRTHG	169	16	80
POL	LVPIHTAELL	17	85	LVPLPIFCILWVYZ	376	13	65
POL	LVVNRPIDW	16	80	VAPPIHTAELLAMC	709	9	45
POL	LVQNSKPG	19	90	FRILQVNRPIDWKVC	608	15	75
X	LVRLPYCAE	18	95	CVMVQENSKRCSQDY	312	10	50
X	LVPRVGAESR	18	90	HL.SL.RGLPYCAESSA	52	18	90
NLC	LVROALCWG	15	75	WOLPRVGAESRGRP	15	18	90
ENV	LVSTPSCFE	19	95	HTALROALCWGELM	52	18	90
NLC	LVSTLVSNAF	19	95	WMLSTPSCFEPSI	237	15	75
ENV	LVSLVPEVO	19	95	LSWL.SLVSNAFYHI	42	10	50
X	LVSTRLPVC	19	95	FSWL.SLVPVOMFV	423	11	55
POL	LVSTFLAOF	17	85	GALVSTRLPVCASF	333	19	95
POL	LVSTKTSFP	18	90	GVQLSTFLAOFISA	50	18	90
ENV	LVSPNRLGIF	15	75	SVLVSTKTSFPMIL	507	17	85
POL	LVTPOSLDS	18	90	TNILLSSNL.SWL.SLDV	739	18	90
ENV	LVTHLTIPO	16	80	GTLVSPNRLGIFPD	405	14	70
POL	LVSKKPNHT	20	100	SSNL.SWL.SLVSNAF	13	17	85
NLC	LVWDFSCFS	18	90	TRILLTIPOSLDSWWT	409	15	75
NLC	LVWHSCLT	17	85	LOSL.TNILLSSNL.SWL	186	18	90
NLC	LVWADIDPY	17	85	FRILLTIHTIPOSID	401	15	75
POL	LVWAGILEYK	20	100	GVFLVWDFSCFSRGN	182	15	75
NLC	LVYREALESP	17	85	LEVLVWDFSCFSRGN	372	11	55
POL	LVYHPILG	16	80	ESLVWDFSCFSRGN	145	14	70
POL	LVYVGLSPL	18	90	POLLVWADIDPYKEF	374	9	45
				LGVLWAGADIDPYKEF	98	17	85
				LHTLVWAGILEYKRET	24	17	85
				ASLVYREALESPERIC	148	18	90
				KULVYHPILGFRHK	34	17	85
				FSVMDVVLGAKSQV	489	16	80
				KIPVAGVLSPLAO	536	18	90
					503	16	80



# HBV DR-Super Motif

Protein	Core Sequence	Core Freq	Core Conservancy (%)	Exemplary Sequence	Position In HBV Poly-Protein	Exemplary Sequence Frequency	Exemplary Sequence Conservancy (%)
POL	MPRLVSSS	17	85	PAAWPLLVSSQSLS	430	8	40
ENV	MDWNSTTH	16	80	PDAMWNSTTHQTL	106	8	40
X	MSITDLEAY	15	75	LSAMSTDL EAYFD	100	9	45
ENV	MMWMPRLY	17	85	IMMMWMPRLVNL	369	9	45
X	WCFSSAGP	19	95	GLPVCASFSAQCAL	57	18	90
POL	WCOHVGIL	17	85	DMKWCORHVGILGFA	618	17	85
POL	VFADAPITG	19	95	LCQAFADAPITGWGL	683	19	95
ENV	VGLSPTWVL	19	95	QMFVGLSPTWVLSVI	344	14	70
POL	VGFLLVNEK	17	85	QOVGFLLVNEKTRL	93	8	40
POL	VHFASPLHV	16	80	PDHVFASPLHVAMR	816	12	60
X	VLCLEPVGA	19	95	ARDMLCLEPVGAEST	12	14	70
POL	VLGAKSVQH	19	95	DDVVLGAKSVQHLES	540	16	80
X	WUIKITLGL	17	85	LPKULIKITLGLSAM	89	11	55
POL	VPLNLSLTN	19	95	KFANPLNLSLTNLS	395	19	95
NJC	VOASKLCLG	16	80	CPTVOASKLCLGLWL	14	15	75
ENV	VHFESWLSIL	16	80	WASVHFESWLSILVPE	328	13	65
POL	VRRAPRPHL	19	95	CSVRAPRPHQLAFS	523	19	95
POL	VSPWTHKV	20	100	NLNVSPWTHKVGNF	45	19	95
NJC	VWRTTPAY	19	95	SFGWRTTPAYRPP	121	18	90
ENV	WVFGLSPTV	19	90	TSFVWVPSALNPADD	764	16	80
NJC	WHHSQLTG	18	90	QLWHHSQLTGHE	99	17	85
ENV	WFLGTSFV	19	95	FVQWFVGLSPTVWLS	342	19	95
POL	WLRTPRVR	16	80	AAMWLRLGTSFVVP	756	14	70
NJC	WRTTPRVR	19	95	FGWRTTPRVRPPN	122	19	95
POL	WKGILYKR	20	100	HTLWKGILYKRETT	149	18	90
POL	WLLGCANWV	16	80	SFPWLLGCANWMLR	748	15	75
ENV	WLSLVSAM	19	95	NLSWLSDVSAEYH	411	17	85
POL	WLSLVPV	20	100	RFSWLSDVSAEYH	332	20	100
ENV	WPKFVAVPL	19	95	RVSMPKFAVPLQSL	390	11	55
POL	YMDVVLGA	18	90	AFSYMDVVLQAKSV	535	18	90
POL	YFALMPLVA	19	95	OCGYFALMPLVACIO	637	19	95
ENV	YOGALPVCP	18	90	LUDYOGALPVCPUP	260	10	50
NJC	YRPPNAPIL	20	100	PPAVRPPNAPILSTL	129	19	95
ENV	YRPPNAPIL	19	95	CRGYRPPNAPILSTL	232	19	95
POL	YSFPIIGF	16	80	LHLYSFPIIGFRKI	490	16	80
POL	YSLNMGV	15	75	RMGYSLNMGVVGGS	588	11	55
ENV	YVPSALNPA	18	90	SFYVPSALNPADDP	765	16	80
ENV	FECQLWYIZ	20	100		382		
ENV	M3TNLSVFN	15	75		12		

148

### Core Sequence

[illegible][illegible]

[illegible]

# ABY DR-Super Motif With Binding Data

Core Sequence	Exemplary Sequence	DR1	DR2w201	DR2w202	DR3	DR4w4	DR4w15	DR5w11	DR5w12	DR6w19	DR7	DR8w2	DR9	DRw53
MPALLVSS	PAAMPPLLVSSSALS	0.0012				0.0300					0.1200			
MCWNSTTH	PCAMCMWNSTTHQTL													
MSTDLFAY	LSAMSTDLFAYFKD													
MMWMPSLY	IMMAMWMPSLYVIL													
VCASSACP	GLPVCASSACAL													
VCORVGL	DMNVCORVGLGFA	0.0120		-0.0026		0.0030		0.2500			0.0018			0.0130
VFADAPFG	LCQVFADAPFGWGL	0.0020				0.9600					0.0013			
VGLSPTWVL	QMFVGLSPTWMLSVI													
VGLLTNVEK	QCVVGLLTNVEKRIL													
VHFAAPLHV	PDRVHFAAPLHVAMR	0.0510	0.0230	0.0008		0.0008	0.5400	0.0008		0.0190	0.0810	0.0035	0.2400	
VLCAPVGA	ARDVLCAPVGAESA													
VLGAKSVOH	DDVVLGAKSVOHLES													
VLIKRTILGL	LPRVLIKRTILGLSAM													
VPNLQSLTN	KFAVPNLQSLTNLLS	0.0180	0.0005	-0.0003		0.1300		0.0043		0.0088	-0.0003		0.0056	
VOASKLCLG	CFTVOASKLCLGMLW													
VRFESMLSL	WMSVRFESMLSLVVF													
VHFAFPIKL	CSVHFAFPIKLQAFS	0.1000	0.1024	0.0770	0.0032	0.0016	-0.2200	0.0008	-0.0013	0.0540	0.0590	0.0250	1.2000	0.0460
VSRPWHKIV	NLNVSRPWHKIVGNF	0.0001		-0.0005	-0.0041	-0.0007		-0.0002			0.0005			0.0009
VMIHTPPAY	SEGVMIHTPPAYRPP	0.0034	0.0110	0.4300	-0.0009	0.0760	0.6300	0.0260	0.0071	0.0002	0.0240	0.2500	0.0800	0.0016
VVPSALNP	TSFVVPSALNPADOD													
WFHSQTLF	QILWFHSQTLFQHE													
WFVGLSPV	FVOMFVGLSPVWLS	0.4700	0.0035	0.0160	-0.0013	0.0130		0.0072	0.0021	0.0190	0.0690	0.0180	0.0410	0.0044
WILGTSFV	ANMWILGTSFVYVP	0.0920	0.0240	0.0061	0.0023	0.0510	0.2500	0.0140	0.3700	0.0250	0.5800	0.2500	0.2700	
WIIHTPPAYR	FGVWIIHTPPAYRPPN													
WKAQILYKR	HTLWKAQILYKRETT													
WILGCAAMW	SFPWLILGCAAMWILR													
WLSLOVSAW	NLSWLSDVSAWFIH	0.1400	0.0003	-0.0005		0.2900		0.0033	0.0022	0.0330	0.0041	0.0150	0.0620	2.4000
WPSLILPEV	RFSWLILPEVQOMF	0.0430		0.0009		-0.0007		0.0002			0.0005			0.0031
WPKFAVPLN	RVSMPKFAVPLNOSL													
YADDDVGLA	AESYADDDVGLAKSV	0.0027		-0.0005	0.0130	2.9000		0.0006			-0.0003			-0.0005
YOCALPVC	QOQYOCALPVCPLUP	0.0062		0.0018		0.0068		0.0023			0.0006			
YCPALMFLYA	LDYOCALMFLYADOD													
YRPANAPIL	PPAYRPPANAPILSTL	0.0056		-0.0005		0.0038		0.0022			0.0024			0.0015
YTHMQLFF	CFYTHMQLFFIFF													
YSHFLLGF	LHLYSHFLLGFRRD	0.0220	0.0340	0.0400	0.0040	0.6800	0.1600	0.0410	0.0310	0.0002	0.0006	0.0610	0.0490	
YSLNMGV	FMGYSLNMGVSGS													
YVPSALNFA	SFYVYVPSALNPADDP													
YFQLWVYZ														
M3TNLSVFN														

148

Table XXa  
IBV DR-3A Motif

Protein	Core Sequence	Core Freq	Core Conservancy (%)	Exemplary Sequence	Position In Poly-Protein	Exemplary Sequence Frequency	Exemplary Sequence Conservancy (%)
ENV	FFPDKQDP	19	95	PLGTFPDKQDPNFG	10	9	95
NJC	FGFETWLEY	15	75	CLTFGEETWLEYVS	136	14	75
POL	FGVGPSSGQ	15	75	FRSGVGPSSGSHID	241	6	75
POL	FLVDKRNPHN	20	100	GGVFLVDKRNPHNTE	360	11	100
POL	IGTDSNMYL	16	80	AKUGTDSNMYLSRK	731	13	80
POL	LEELPPLA	16	90	AGPLLEELPPLADEQ	18	13	90
POL	LPDKGKRP	20	100	TKVLPDKGKRPYP	120	20	100
POL	LSLDVSAAF	19	95	LSWLSLDVSAAFYHI	412	11	95
POL	LVDFSCFS	20	100	ESPLVDFSCFSNIN	374	9	100
NJC	LYTEALESP	17	85	ASALYTEALESPFHC	34	17	85
NJC	MDPYKEF	17	85	LMGMDPYKEFGAS	27	9	85
POL	VAEDLNIGN	20	100	NHRVAEDLNIGNLV	34	17	100
POL	VFADATPTG	19	95	LCOVFADATPTGWGL	683	19	95
ENV	VILDYQGM	19	95	FLVILDYQGMPLVC	256	18	95
POL	YMDDVWLGA	18	90	AFSYMDDVWLGAHSV	535	18	90

TABLE XXb  
VDR 3A Motif

Core Sequence	Exemplary Sequence	DR1	DR2w231	DR2w232	DR3	DR4w4	DR4w15	DR5w11	DR5w12	DR6w1	DR7	DR8w2	DR9	DRw53
FFPDHQLDP	FLGFFPDHQLDPARQ													
FGRETMLEY	CLTFGRETMLEYLVS													
FGVPSQSSG	FFSGVPSQSSGHD				0.0790									
FLVDKNPIN	GGVFLVDKNPINITE				0.0022									
IGTDSNVVL	AKLIGTDSNVLSHK				-0.0017									
LEELUPILA	AGPLEELUPILADEG													
LPLDKGKP	TKYLPDKGKPYTP													
LSLDVSAF	LSWLSLDVSAFETH													
LWDFSCFS	ESRLWDFSCFSRGN	0.0007	0.0074	-0.0010	2.6000			-0.0004		0.4000	-0.0014	0.0029		
LYHEALESP	ASALYHEALESPENC													
MDDFYKEF	LWGMDDFYKEFGAS				0.1400									
VAEIDLIGN	NRRVAEIDLIGNLW	0.0020			0.0170	0.9600				0.0013				
VFADATPTG	LCOVFADATPTGWMQ													
VLLDYOGML	FLWVLLDYOGMLPVC				0.0130	2.9000		0.0006		-0.0003				
YMDVVLGA	AFSYMDVVLGAQKSV	0.0027		-0.0005										-0.0005

15

Table XXc  
HBV DR-3B Motif

Protein	Core Sequence	Core Freq	Core Conservancy (%)	Exemplary Sequence	Position In HBV Poly-Protein	Exemplary Sequence Frequency	Exemplary Sequence
X	ALLSLRLP	18	90	DIQALSLRLPVCA	48	18	90.00
POL	FSPTTKAEL	19	95	AFESPTYKALCKQ	555	11	55.00
POL	IPMTIKVGN	20	100	MSIPMTIKVGNFTG	47	20	100.00
POL	LTVNEKRRL	17	85	VGFLTVNEKRRLKU	96	12	60.00
X	VGKESRRP	19	95	LPVGVKESRRPVSQ	18	7	35.00
POL	VLSRKLYTS	18	90	DNSVWLSSRKLYTSFW	737	17	85.00

Table XXd  
HBV DR-3B Motif With Binding Information

Core Sequence	Exemplary Sequence	DR1	DR2w201	DR2w202	DR3	DR4w4	DR4w15	DR5w11	DR5w12	DR6w19	DR7	DR8w2	DR9	DRw53
AHSLRGLP FSPTKAKL IPWTHKQGN LTYNEKRL VGAESTGRP VLSIKYIS	DHGAHSLRGLPVCA AFTSPTYKALCKO NWSIPWTHKQGNFTG VGPLTYNEKRLKU LRFVGAESTGRPVSG DMSVLSIKYISFTPW				0.0035									
		0.0006	0.0022	0.0047	2.2000 -0.0017			0.0030		0.0009	-0.0014	0.0092		

HBV DR-3B Motif With Binding Information  
Table XXd



TABLE XXI. Population coverage with combined HLA Supertypes

HLA-SUPERTYPES	PHENOTYPIC FREQUENCY					
	Caucasian	North American Black	Japanese	Chinese	Hispanic	Average
<b>a. Individual Supertypes</b>						
A2	45.8	39.0	42.4	45.9	43.0	43.2
A3	37.5	42.1	45.8	52.7	43.1	44.2
B7	38.6	52.7	48.8	35.5	47.1	44.7
A1	47.1	16.1	21.8	14.7	26.3	25.2
A24	23.9	38.9	58.6	40.1	38.3	40.0
B44	43.0	21.2	42.9	39.1	39.0	37.0
B27	28.4	26.1	13.3	13.9	35.3	23.4
B62	12.6	4.8	36.5	25.4	11.1	18.1
B58	10.0	25.1	1.6	9.0	5.9	10.3
<b>b. Combined Supertypes</b>						
A2, A3, B7	83.0	86.1	87.5	88.4	86.3	86.2
A2, A3, B7, A24, B44, A1	99.5	98.1	100.0	99.5	99.4	99.3
A2, A3, B7, A24, B44, A1, B27, B62, B58	99.9	99.6	100.0	99.8	99.9	99.8

SF 184895 v1

Table XXII HBV ANALOGS

AA	Sequence	Fixed Nomen.	A1 Moll	A2 Super Moll	A3 Super Moll	A24 Moll	B7 Super Moll	Anchor Fixer	Analog
10	CILLCLFL	VM2.V9	N	Y	N	N	N	No	A
9	RMGTGVFLV	VM2.V9	N	Y	N	N	N	1	A
9	LMPEVQWV	VM2.V9	N	Y	N	N	N	1	A
9	RLTGAVFLV	VL2.V9	N	Y	N	N	N	1	A
9	GLCOVFAOV	L2.AV9	N	Y	N	N	N	1	A
9	WLLRGTSFV	IL2.V9	N	Y	N	N	N	1	A
9	NLGNLNVSV	L2.IV9	N	Y	N	N	N	1	A
9	YLPNALNPV	VL2.AV9	N	Y	N	N	N	1	A
9	GLWRTIPV	VL2.AV9	N	Y	N	N	N	1	A
9	RLSWPKFAV	VL2.V9	N	Y	N	N	N	1	A
9	ILGLGFAV	VL2.AV9	N	Y	N	N	N	1	A
9	RMLTPQSV	IM2.LV9	N	Y	N	N	N	1	A
9	SLDSWWTSV	L2.LV9	N	Y	N	N	N	1	A
10	FMILLCLFL	IM2.L10	N	Y	N	N	N	1	A
10	LMLOAGFELV	VM2.LV10	N	Y	N	N	N	1	A
10	SMLSPLPLV	IM2.LV10	N	Y	N	N	N	1	A
10	LMLDYOGMV	VM2.LV10	N	Y	N	N	N	1	A
10	FLGLSPTVMV	VL2.LV10	N	Y	N	N	N	1	A
8	FPAAMPHL		N	N	N	N	N		A
8	HPFAMPHL		N	N	N	N	N		A
8	HPAAMPHI		N	N	N	N	N		A
8	FMFSPTYK		N	N	N	N	N		A
8	FVFSPTYK		N	N	N	N	N		A
9	FLITRLITV	L2.IV9	N	Y	N	N	N	1	A
9	ALMPLYACV	L2.IV9	N	Y	N	N	N	1	A
9	LLOFTSAV	L2.IV9	N	Y	N	N	N	1	A
9	LPPVQWV	VL2.V9	N	Y	N	N	N	1	A
9	FLAOFTSV	L2.AV9	N	Y	N	N	N	1	A
9	KLHLYSHPV	L2.IV9	N	Y	N	N	N	No	A
9	KLFLYSHPI		N	Y	N	N	N	1	A
9	LSSNLISWV	L2.LV9	N	Y	N	N	N	1	A
9	FLSLGIHV	L2.LV9	N	Y	N	N	N	1	A
9	NMWWYWGSPV	M2.LV9	N	Y	N	N	N	1	A
9	VLOAGFELV	L2.LV9	N	Y	N	N	N	1	A
9	PLPIFEV	L2.LV9	N	Y	N	N	N	No	A
9	FLPIFECL	L2.LV9	N	Y	N	N	N	No	A
9	VLLDYOGMV		N	Y	N	N	N	1	A
9	YMFEDVLGA	L2.AV9	N	Y	N	N	N		A
9	GLGWSPOV		N	N	N	N	N		A
9	FPAAMPHLL		N	N	N	N	N		A
9	HPFAMPHLL		N	N	N	N	N		A
9	HPAAMPHLI		N	N	N	N	N		A
9	FPVCAFFSSA		N	N	N	N	N		A
9	LPFCAFFSSA		N	N	N	N	N		A
9	LPVCAFFSSI		N	N	N	N	N		A

AA Sequence

Fixed  
Normen.

A1 A2 A3  
Moll Super Moll Super Moll Super

A24  
Moll

B7  
Super Moll

Anchor  
Fixer

Analogs

# HBV ANALOGS

9	FPALMPLYA	N	N	N	N	N	N	Y	A
9	YPALMPLYI	N	N	N	N	N	N	Y	A
9	FPSPGLGL	N	N	N	N	N	N	Y	A
9	DPPRGLGL	N	N	N	N	N	N	Y	A
9	DPSRGLGI	N	N	N	N	N	N	Y	A
9	SMICSVRR	N	N	N	N	N	N	N	A
9	SVICSVRR	N	N	N	N	N	N	N	A
9	KVGNFTGUK	N	N	N	N	N	N	N	A
9	KVGNFTGUT	N	N	N	N	N	N	N	A
9	VFFSQFSR	N	N	N	N	N	N	N	A
9	SVNRPIDWK	N	N	N	N	N	N	N	A
9	TLWKAGILK	N	N	N	N	N	N	N	A
9	TLWKAGILK	N	N	N	N	N	N	N	A
9	TLWKAGILK	N	N	N	N	N	N	N	A
9	TLWKAGILK	N	N	N	N	N	N	N	A
9	TLWKAGILY	N	N	N	N	N	N	N	A
9	TVWKAGILY	N	N	N	N	N	N	N	A
9	RMVLTILWK	N	N	N	N	N	N	N	A
9	RVYLHTLWK	N	N	N	N	N	N	N	A
9	AMTFSPTYK	N	N	N	N	N	N	N	A
9	SVRRAPFR	N	N	N	N	N	N	N	A
9	SVRRAPFK	N	N	N	N	N	N	N	A
9	SAIXSVRR	N	N	N	N	N	N	N	A
9	LPVXAFSSA	N	N	N	N	N	N	N	A
10	FLAOFISAV	N	N	N	N	N	N	N	A
10	YLFITLWKAGI	N	N	N	N	N	N	N	A
10	YLTITLWKAGI	N	N	N	N	N	N	N	A
10	LLPYOGMLPV	N	N	N	N	N	N	N	A
10	LLYOGMLPV	N	N	N	N	N	N	N	A
10	LLVLOAGFV	N	N	N	N	N	N	N	A
10	ILLCLRELY	N	N	N	N	N	N	N	A
10	FPCLAFSYM	N	N	N	N	N	N	N	A
10	FPCLAFSYI	N	N	N	N	N	N	N	A
10	FPARVTGVF	N	N	N	N	N	N	N	A
10	TPFRVTGVF	N	N	N	N	N	N	N	A
10	TPARVTGVI	N	N	N	N	N	N	N	A
10	FPCLARFTSA	N	N	N	N	N	N	N	A
10	GPCALRFTSI	N	N	N	N	N	N	N	A
10	FPAMPHLLV	N	N	N	N	N	N	N	A
10	HPFAMPHLLV	N	N	N	N	N	N	N	A
10	HPAAMPHLLI	N	N	N	N	N	N	N	A
10	OMFTFSPTYK	N	N	N	N	N	N	N	A
10	OVFTFSPTYK	N	N	N	N	N	N	N	A
10	TMWKAGILYK	N	N	N	N	N	N	N	A
10	TVWKAGILYK	N	N	N	N	N	N	N	A

L2.LV10  
L2.LV10

1  
No  
No  
No  
1

# HBV ANALOGS

AA	Sequence	Fixed Nomen.	A1 Mottl	A2 Super Mottl	A3 Super Mottl	A24 Mottl	B7 Super Mottl	Anchor Fixer	Analog
10	VMGGVELVDK		N	N	Y	N	N		A
10	VWGGVELVDK		N	N	Y	N	N		A
10	SMLPETTVRR		N	N	Y	N	N		A
10	SVLPETTVRR		N	N	Y	N	N		A
10	TMPEITTVRR		N	N	Y	N	N		A
10	TVPETTVRR		N	N	Y	N	N		A
10	HTLWKAGILR		N	N	Y	N	N		A
10	HTLWKAGILK		Y	N	Y	N	N		A
10	HMLWKAGILY		N	N	Y	N	N		A
10	HVLWKAGILY		N	N	Y	N	N		A
10	GMDNSVLSRI		N	N	Y	N	N		A
10	GVDNSVLSRI		N	N	Y	N	N		A
10	GTFSNVLSRI		N	N	Y	N	N		A
10	YMFVVLGAK		N	N	Y	N	N		A
10	MMWYWGPSLK		N	N	Y	N	N		A
10	MMWYWGPSLR		N	N	Y	N	N		A
9	ILLXLIFL		N	Y	N	N	N		A
9	LLXLIFLL		N	Y	N	N	N		A
9	LLXLIFLV		N	Y	N	N	N		A
9	PLPIFEFL		N	Y	N	N	N		A
9	ALMPLYAXI		N	Y	N	N	N		A
9	GLXQVFADA		N	N	Y	N	N		A
9	HISXLTGFR		N	N	Y	N	N		A
9	FVLGGXRIK		N	Y	N	N	N		A
10	FILLXLIFL		N	Y	N	N	N		A
10	ILLXLIFLL		N	Y	N	N	N		A
10	LLXLIFLV		N	Y	N	N	N		A
10	LLPIFEFLW		N	Y	N	N	N		A
10	QLWFIHSXL		N	Y	N	N	N		A
10	LLGXANWIL		N	N	Y	N	N		A
10	TSAIXSVRR		N	N	N	N	N		A
10	GYRWMLRRF		N	N	N	N	N		A
10	GPXALRFTA		N	N	N	N	N		A
10	FPHXLAFTYM		N	N	Y	N	N		A
10	HMLWKAGILYK		N	N	Y	N	N		A
11	HVLWKAGILYK		N	N	Y	N	N		A
11	SMLPETTVRRR		N	N	Y	N	N		A
11	SVLPETTVRRR		N	N	Y	N	N		A
11	GMDNSVLSRK		N	N	Y	N	N		A
11	GVDNSVLSRK		N	N	Y	N	N		A
11	GTFSNVLSRK		N	N	Y	N	N		A
8	MPLSYOHI		N	N	N	N	Y		A
8	LPIFFCLI		N	N	N	N	Y		A
8	SPFLAOL		N	N	N	N	Y		A
8	YPALMPLI		N	N	N	N	Y		A

## HBV ANALOGS

AA	Sequence	Fixed Nomen.	A1 Mottl	A2 Super Mottl	A3 Super Mottl	A24 Mottl	B7 Super Mottl	1° Anchor Fixer	Analog
8	VPSALNPI		N	N	N	N	Y		A
9	LPIFFCLWI		N	N	N	N	Y		A
9	LPIHTAELI		N	N	N	N	Y		A
10	VPRQWVRGI		N	N	N	N	Y		A
11	NPLGFFPDHDI		N	N	N	N	Y		A
11	LPIHTAELLAI		N	N	N	N	Y		A
9	FLPSYFPPSA		N	Y	N	N	N		A
10	YUHTLWKAGV		N	Y	N	N	N		A
11	STLPETYVVR		N	Y	N	N	N		A
9	YMDOWMGV		N	Y	N	N	Y		A
9	FPIPPSSWAF		N	N	N	N	Y		A
9	IPITSSWAF		N	N	N	N	Y		A
9	IPILSSWAF		N	N	N	N	Y		A
9	FPVCLAFSY		N	N	N	N	Y		A
9	FPHCLAFAY		N	N	N	N	Y		A
9	FPIICLAFSL		N	N	N	N	Y		A
9	IPIPMSWAF		N	N	N	N	Y		A
9	FPHCLAFAL		N	Y	N	N	N		A
10	FLPSZFPPSV		N	Y	N	N	Y		A
10	FLPSZFPPSV		N	N	N	N	Y		A
9	IPEPSSWAF		N	N	N	N	Y		A
9	IPIPSSWAI		N	N	N	N	Y		A
9	FPFCLAFSY		N	N	N	N	Y		A
9	FPHCLAFSI		N	N	N	N	Y		A
9	FPHCLAFSA		N	Y	N	N	N		A
10	FOPSDYFPPSV		N	Y	N	N	N		A
9	YLTTRILTI		N	Y	N	N	N		A
9	FLYTRILTI		N	Y	N	N	N		A
9	FLTYILTI		N	Y	N	N	N		A
9	FLTRILYI		N	N	N	N	N		A
11	FLPSDFPPSVR		N	N	N	N	N		A
9	FLPSDFPPS		N	N	N	N	N		A
8	FLPSDFPP		N	Y	N	N	N		A
10	FLPSDFPPSI		N	Y	N	N	N		A
10	FLPSDYFPPSV		N	N	N	N	N		A
12	YSEFLPSDFPPSV		N	N	N	N	N		A
10	YMKGLKFRQL		N	Y	N	N	N		A
9	NMGJLYRQL		N	N	N	N	N		A
10	FLPS(X)YFPPSV		N	N	N	N	N		A
10	FLPSD(X)FPPSV		N	N	N	N	N		A
11	FLPSDILPSVR		N	N	N	N	N		A
12	FLPSDFPPSVRD		N	N	N	N	N		A
12	LSFLPSDFPPSV		N	N	N	N	N		A
11	SFLPSDFPPSV		N	N	N	N	N		A
8	PSDFPPSV		N	N	N	N	N		A

L2.V110

M2.AV9

L2.FV5.VA9  
L2.IV10

1

Rev3  
1

Rev

Nb

Rev  
Nb

Nb

## HBV ANALOGS

AA	Sequence	Fixed Noncon.	A1 Motif	A2 Super Motif	A3 Super Motif	A24 Motif	B7 Super Motif	1° Anchor Fixer	Analog
9	FLMSYFPPSV	L2.FY5.V9	N	Y	N	N	N	Nb	A
9	FLPSYFPPSV		N	Y	N	N	N	3	A
10	FLMSDYFPPSV		N	Y	N	N	N	Nb	A
11	CILLCLIFLL		N	Y	N	N	N	Rev	A
10	FLPNDFFPSA	L2.SN4.VA10	N	Y	N	N	N	Rev	A
10	FLPNDFFPSA	L2.SD4.VA10	N	Y	N	N	N	Nb	A
10	FLPNDFFPSV		N	Y	N	N	N	Rev	A
10	FLPSDFFPSA	L2.VA10	N	Y	N	N	N	Nb	A
10	FLPSDFFPSV		N	Y	N	N	N	Nb	A
10	FLPADFFPSV		N	Y	N	N	N	Rev	A
10	FLPADFFPSI	L2.SA4.VI10	N	Y	N	N	N	Rev	A
10	FLPSDAFPPSV	L2.SV4.VI10	N	Y	N	N	N	Nb	A
10	FLPSAEFPPSV		N	Y	N	N	N	Nb	A
10	FLPSDFAPSV		N	Y	N	N	N	Nb	A
10	FLPSDFFASV		N	Y	N	N	N	Nb	A
10	FLPSDFFPAV		N	Y	N	N	N	Rev	A
10	FLASDFFPSV	LA2.V10	N	Y	N	N	N	Nb	A
10	FAPSDFFPPSV		N	Y	N	N	N	Nb	A
10	ALPSDFFPSV		N	Y	N	N	N	1	A
10	FMPSDFFPPSV	LM2.V10	N	Y	N	N	N	Nb	A
10	FLKSDFFPSV		N	Y	N	N	N	Nb	A
10	FLPSEFFPSV		N	Y	N	N	N	Nb	A
10	FLPSDFYPSV		N	Y	N	N	N	Nb	A
10	FLPSDFFKSV		N	Y	N	N	N	Nb	A
10	FLPSDFFPKV		N	Y	N	N	N	Nb	A
10	FLPSDFFPSV(CONH2)		N	N	N	N	N	Nb	Amidated
10	VLEYLVSEFGV(NH2)		N	N	N	N	N	Nb	Amidated
17	ATVELLSFLPSDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
16	TVELLSFLPSDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
15	VELLSFLPSDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
14	ELLSFLPSDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
13	LLSFLPSDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
12	LSFLPSDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
11	SFLPSDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
10	FLPSDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
9	LPDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
8	PSDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
9	FLPSDFFPS.NH2		N	N	N	N	N	Nb	Amidated
8	FLPSDFFP.NH2		N	N	N	N	N	Nb	Amidated
7	FLPSDFF.NH2		N	N	N	N	N	Nb	Amidated
10	ALPSDFFPPSV.NH2		N	N	N	N	N	Nb	Amidated
10	SLNFLGGTIV(NH2)		N	N	N	N	N	Nb	Amidated

# HBV ANALOGS

AA	Sequence	Fixed Normen.	A1 Mottl	A2 Super Mottl	A3 Super Mottl	A24 Mottl	B7 Super Mottl	1° Anchor Fixer	Analog
11	FLPSDEFPSPVR-NH2		N	N	N	N	N		Amidated
9	ALFKDWIEL		N	Y	N	N	N		A
9	VLGGSFHKL		N	Y	N	N	N		A
9	KIKESFRKL		N	Y	N	N	N		A
9	ALMPLYASI		N	Y	N	N	N		A
9	FLSKOYLNL		N	Y	N	N	N		A
9	LLGSAANIWI		N	Y	N	N	N		A
9	NLNNLNWSI		N	Y	N	N	N		A
9	IKKSEQPV		N	Y	N	N	N		A
9	ALSLIVNLL		N	Y	N	N	N		A
9	RLPNTPRSV		N	Y	N	N	N		A
237									

Table XXIII: Immunogenicity of HBV-derived peptides

Supermotif	Peptide	Sequence	Protein	XRN	Immunogenicity		
					primary	transgenic	patients overall <sup>a</sup>
A2 supermotif	924.07	FLPSDFPPSV	HBV core 18	5	10/10	6/6	25/32 <sup>a</sup>
	1069.06	LLVPFVQWFFV	HBV env 338	5	3/4	6/9	+
	1147.13	FLAQFTSAI	HBV pol 513	5		0/3	unk
	1090.77	YMDDVVLGV	HBV pol 538	5		9/9	+
	777.03	FLTRLTLI	HBV env 183	4			14/23 <sup>a</sup>
	927.15	ALMPLVACI	HBV pol 642	4	10/12	3/5	2/15 <sup>a</sup>
	1013.01	WLSLVFV	HBV env 335	4	2/6	5/9	23/29 <sup>a</sup>
	1069.05	LLAQFTSAI	HBV pol 504	4	0/4	0/5	+
	1132.01	LVPFVQWFFV	HBV env 339	4	0/3	0/4	unk
	1147.14	VLLDYQGMLPV	HBV env 259	4	4/4	6/6	unk
	927.41	LSSNLSWL	HBV pol 992	3	0/4	0/3	+
	927.42	NLSWLSLDV	HBV pol 411	3		2/8	+
	927.46	KLHLYSHPI	HBV pol 489	3	0/4	4/6	+
	1069.07	FLAQFTSA	HBV pol 503	3	1/2	0/3	+
A2 supermotif	1168.02	GLSRVVARL	HBV pol 455	3			9/13 <sup>a</sup>
	927.11	FLSLGIHL	HBV pol 562	2	15/22	12/13	9/15 <sup>a</sup>
	927.47	HLYSHPIIL	HBV pol 1076	2		10/14	+
	1039.03	MMWYWGPSL	HBV env 360	2	3/4	0/4	+
	1069.12	YLHTLWKAGV	HBV pol 147	2	2/4		+
	1137.02	LLDYQGMLPV	HBV env 260	2	1/2	0/4	+
	1142.07	GILGWSPQA	HBV env 62	2	3/4	5/6	+
	1.0573	ILRGTSFVYV	HBV pol 773	1			3/7 <sup>b</sup>
	1013.14	VLOAGFFLL	HBV env 177	1	0/4	5/12	+
	1069.10	LLPIFFCLWV	HBV env 378	1	3/3	0/4	2/5 <sup>c</sup>
	1069.13	PLLPFFCL	HBV env 377	1	0/4	7/12	+
	1090.06	LLVLQAGFFL	HBV env 175	1	1/5	0/4	+
	1090.12	YLVFSGVWI	HBV nuc 118	1	9/9		+
	1.0518	GLSPTVWLSV	HBV env 338	1			3/9 <sup>c</sup>
A3 supermotif	1090.14	YMDDVVLGA	HBV pol 538	1	2/7	2/5	2/7 <sup>b</sup>
	1147.16	HTLWKAGILYK	HBV POL 149	5	0/6	3/3	1/22
	1083.01	STLPEITVVR	HBV core 141	4	3/5	6/6	8/32
	1150.51	GSTHVSWP	HBV pol 398	4		3/6	+
	1.0219	FVLGCCRHK	HBV adr "X" 1550	3	0/4		unk
	1069.16	NVSIPWTHK	HBV pol 47	3	0/8	0/3	1/21



A3 supernotif	1069.20	LVVDFEQFSR	HBV pol 388	3	0/4	6/6	1/22	+
	1090.10	QAFTEPTPYK	HBV pol 665	3	3/6	0/3	3/21	+
	1090.11	SAICSVVRR	HBV pol 531	3	1/4		2/22	+
A3 supernotif	1069.15	TLWKAGILYK	HBV pol 150	2	3/8	0/3	5/28	+
	1142.05	KVGNFTGLY	HBV adr POL 629	2		0/3	2/22	+
B7 supernotif	1147.05	FPHCLAFSYM	HBV POL 530	5	1/3		0/12	+
	988.05	LPSEFFPSV	HBV core 19-27	4			2/16	+
	1145.04	IPIPSSWAF	HBV ENV 313	4	0/4		1/12	+
	1147.02	HPAAMPHL	HBV POL 429	4	0/5		0/12	unk
	1147.06	LPVCAFSFA	HBV X 58	4	1/4			+
	1147.08	YPALMPLYA	HBV POL 640	4			0/12	unk
	1145.08	FPHCLAFSYM	HBV POL 541	3	0/4			unk
B7 supernotif	1147.04	TPARVTGGVF	HBV POL 354	2			2/12	+

Immunogenicity evaluation derived from primary cultures, acute patients (a-Bertoni et al, J Clin Invest 100:503, b-Rehermann et al., J. Clin. Invest 97:1655, c-Nayersina et al., J Immunol 150:4639) or transgenic mice. A positive assessment (+) is assigned when responders have been noted in one of these systems. Unk=unknown

Table XXIV. MHC-peptide binding assays: cell lines and radiolabeled ligands.

## A. Class I binding assays

Species	Antigen	Allele	Cell line	Radiolabeled peptide		Notes
				Source	Sequence	IC50 nM
Human	A1	A*0101	Steinlin	Hu. J chan 102-110	YTAVVPLVY	no NEN in P1 cocktail
	A2	A*0201	JY	HBVc 18-27 F6->Y	FLPSDYFPSV	
	A2	A*0202	P815 (transfected)	HBVc 18-27 F6->Y	FLPSDYFPSV	"
	A2	A*0203	FUN	HBVc 18-27 F6->Y	FLPSDYFPSV	"
	A2	A*0206	CLA	HBVc 18-27 F6->Y	FLPSDYFPSV	"
	A2	A*0207	21.221 (transfected)	HBVc 18-27 F6->Y	FLPSDYFPSV	"
	A3		GM3107	non-natural (A3CON1)	KVFPYALINK	"
	A11		BVR	non-natural (A3CON1)	KVFPYALINK	"
	A24	A*2402	KAS116	non-natural (A24CON1)	AYIDNYNKF	"
	A31	A*3101	SPACH	non-natural (A3CON1)	KVFPYALINK	"
	A33	A*3301	LWAGS	non-natural (A3CON1)	KVFPYALINK	"
	A28/68	A*6801	C1R	HBVc 141-151 T7->Y	STLPETYVVR	"
	A28/68	A*6802	AMAI	HBV pol 646-654 C4->A	FTQAGYPAL	"
	B7	B*0702	GM3107	A2 sigal seq. 5-13 (L7->Y)	APRTLVELL	"
	B8	B*0801	Steinlin	[Vgp 586-593 Y1->F, Q5->R 60s	FLKDYQLL	"
	B27	B*2705	LG2		FRYNGLIHR	"
	B35	B*3501	C1R, BVR	non-natural (B35CON2)	FPEKYAAAF	"
	B35	B*3502	TISI	non-natural (B35CON2)	FPEKYAAAF	"
	B35	B*3503	EHM	non-natural (B35CON2)	FPEKYAAAF	"
	B44	B*4403	PITOUT	EF-1 G6->Y	AEMGKYSFY	"
Mouse	B51		KAS116	non-natural (B35CON2)	FPEKYAAAF	"
	B53	B*5301	AMAI	non-natural (B35CON2)	FPEKYAAAF	"
	B54	B*5401	KT3	non-natural (B35CON2)	FPEKYAAAF	"
	Cw4	Cw*0401	C1R	non-natural (C4CON1)	QYDDAVYKL	"
	Cw6	Cw*0602	21.221 transfecte	non-natural (C6CON1)	YRHDGNNVL	"
	Cw7	Cw*0702	21.221 transfecte	non-natural (C6CON1)	YRHDGNNVL	"
	D <sup>b</sup>		EL4	Adenovirus E1A P7->Y	SGPSNTYPEI	"
	K <sup>b</sup>		EL4	VSV NP 52-59	RGYVFQGL	"
	D <sup>d</sup>		P815	HIV-III <sub>B</sub> ENV G4->Y	RGPYRAFTI	"
	K <sup>d</sup>		P815	non-natural (KdCON1)	KFNPMKTYI	"
L <sup>d</sup>			P815	HBVs 28-39	IPQSLDSYWTSL	"

# B. Class II binding assays

Species	Antigen	Allele	Cell line	Radiolabeled peptide			Notes
				Source	Sequence	IC50 nM	
Human	DR1	DRB1*0101	LG2	HA Y307-319	YPKYVKQNTLKLAT		optimal assay pH is 4.;
	DR2	DRB1*1501	L466.1	MBP 88-102Y	VVHFFKNIVTPRTPPY		
	DR2	DRB1*1601	L242.5	non-natural (760.16)	YAAFAAAKTAFAFA		
	DR3	DRB1*0301	MAT	MT 65kD Y3-13	YKTAFDDEARR		
	DR4w4	DRB1*0401	Preiss	non-natural (717.01)	YARFQSQTTLKQKT		
	DR4w10	DRB1*0402	YAR	non-natural (717.10)	YARFQRQTTLKAAA		
	DR4w14	DRB1*0404	BIN 40	non-natural (717.01)	YARFQSQTTLKQKT		
	DR4w15	DRB1*0405	KT3	non-natural (717.01)	YARFQSQTTLKQKT		
	DR7	DRB1*0701	Pitout	Tet. tox. 830-843	QYIKANSKFIGITE		
	DR8	DRB1*0802	OLL	Tet. tox. 830-843	QYIKANSKFIGITE		
	DR8	DRB1*0803	LUY	Tet. tox. 830-843	QYIKANSKFIGITE		
	DR9	DRB1*0901	IID	Tet. tox. 830-843	QYIKANSKFIGITE		
	DR11	DRB1*1101	Sweig	Tet. tox. 830-843	QYIKANSKFIGITE		
	DR12	DRB1*1201	Heiluf	unknown eluted peptide	EALIHQLKINPYVLS		
	DR13	DRB1*1302	H0301	Tet. tox. 830-843 S->A	QYIKANAKFIGITE		
	DR51	DRB5*0101	3M3107 or L416.;	Tet. tox. 830-843	QYIKANAKFIGITE		
	DR51	DRB5*0201	L255.1	HA 307-319	PKYVKQNTLKLAT		
	DR52	DRB3*0101	MAT	Tet. tox. 830-843	NGQIGNDPNRDIL		
	DR53	DRB4*0101	L257.6	non-natural (717.01)	YARFQSQTTLKQKT		
Mouse	DQ3.1	QA1*0301/DQB1*0301	PF	non-natural (ROIV)	AHAHAHAHAHAHAHA		no NEM in PI mix
	IA <sup>b</sup>		DB27.4	non-natural (ROIV)	AHAHAHAHAHAHAHA		optimal assay pH is 5.;
	IA <sup>d</sup>		A20	non-natural (ROIV)	AHAHAHAHAHAHAHA		
	IA <sup>k</sup>		CH-12	HEL 46-61	YNTDGSTDYGLQINSR		optimal assay pH is 5.;
	IA <sup>s</sup>		LS102.9	non-natural (ROIV)	AHAHAHAHAHAHAHA		
	IA <sup>u</sup>		91.7	non-natural (ROIV)	AHAHAHAHAHAHAHA		
	IE <sup>d</sup>		A20	Lambda repressor 12-26	YLEDARRRKKAIYEKKK		optimal assay pH is 5.;
	IE <sup>k</sup>		CH-12	Lambda repressor 12-26	YLEDARRRKKAIYEKKK		optimal assay pH is 5.;

Table XXV. Monoclonal antibodies used in MHC purifi

Monoclonal antibody	Specificity
W6/32	HLA-class I
B123.2	HLA-B and C
IVD12	HLA-DQ
LB3.1	HLA-DR
M1/42	H-2 class I
28-14-8S	H-2 D <sup>b</sup> and L <sup>d</sup>
34-5-8S	H-2 D <sup>d</sup>
B8-24-3	H-2 K <sup>b</sup>
SF1-1.1.1	H-2 K <sup>d</sup>
Y-3	H-2 K <sup>b</sup>
10.3.6	H-2 IA <sup>k</sup>
14.4.4	H-2 IE <sup>d</sup> , IE <sup>K</sup>
MKD6	H-2 IA <sup>d</sup>
Y3JP	H-2 IA <sup>b</sup> , IA <sup>s</sup> , IA <sup>u</sup>

**Table XXVI: in vitro binding of conserved HBV-derived peptides to HLA-A2-supertype alleles.**

Peptide	AA	Molecule	1st Pos	Sequence	Cons. <sup>1</sup>	A2-supertype binding capacity (IC50 nM)					Alleles <sup>2</sup> bound
						A*0201	A*0202	A*0203	A*0206	A*6802	
924 07	10	Core	18	FLRSPDFPSV	95	2.5	2.1	6.0	3.0	36	5
1069 06	10	ENV	349	LLVPFQWFV	95	7.5	11	5.9	13	286	5
1147.13	10	POL	524	FLLAQFTSAI	95	24	134	1.4	34	455	5
1013 0102	9	ENV	346	WLSLLVPFV	100	4.6	113	1.4	10	1290	4
777 03	9	ENV	183	FLTLRLTI	80	9.8	100	1.3	19	3	4
927 15	9	POL	653	ALMPL YACI	95	10	126	3.0	160	851	4
1069 05	9	POL	525	LLAQFTSAI	95	50	16	3.0	1538	51	4
1132 01	9	ENV	350	LVPFQWFV	95	119	287	2083	463	14	4
1147 14	11	ENV	259	VLLDYQGM L P V	90	8.6	20	2.0	13	2353	4
1090 77	9	POL	538 (a)	YMDVYVLGV	90	5.1	90	6.7	71	1905	4
1069 071	9	POL	524	FLLAQFTSA	95	6.0	1654	9.1	39	870	3
927 46	9	POL	500	KLHL YSHPI	95	72	126	3.7	627	26667	3
927 42	9	POL	422	NLSWLSDV	90	77	843	16	2313	404	3
1168 02	9	POL	455	GLSRVVARL	90	79	391	18	12333	-	3
927 41	9	POL	418	LSSNLSWL	90	455	55	2.6	1370	4000	3
1039 031	9	ENV	360	MMWYWGPSL	85	5.6	5375	833	112	3636	2
927.11	9	POL	573	FLLSLGHL	95	7.7	4300	1000	34	11429	2
1142.07	9	ENV	73	GLLGWSPQA	85	13	14333	286	1429	-	2
927 47	9	POL	502	HL YSHPIIL	80	23	14333	11	2176	755	2
1137.02	10	ENV	271	LDDYQGM L P V	90	51	-	500	552	-	2
1069 09	9	ENV	270	VLLDYQGM L	95	114	-	476	4111	-	2
1069.14	10	NUC	168	ILSTLPETTV	100	238	506	130	1194	5970	2
1069.11	10	POL	147	YLHTLWKAGI	100	313	8600	18	4000	1250	2
1142 01	9	NUC	129	LLWFHISCL	90	385	21500	238	1194	4082	2
1090 12	9	NUC	147	YLVSEGVWI	90	13	-	-	-	-	1
1 0518	10	ENV	359	GLSPTVWLSV	75	18	-	-	-	-	1
1013 1402	9	ENV	177	VLQAQGFLL	95	33	2389	3704	1947	6349	1
1069.13	9	ENV	388	PLPIEFCL	100	77	-	5556	3364	8511	1
1069.10	10	ENV	389	LLPIEFCLWV	100	156	5375	667	5000	-	1
1090.06	10	ENV	175	LLVLQAQGFLL	90	161	1162	2222	2467	3636	1
1 0895	10	ENV	248	FILLCLIFL	80	179	-	-	-	-	1
927.24	9	POL	770	WILRGTSFV	80	185	-	-	-	-	1
1090.14	9	POL	538	YMDDVVLGA	90	200	-	4167	-	-	1
3 0205	10	ENV	171	FLGRLVLQA	75	263	-	-	-	-	1
1069 08	10	ENV	260	ILLCLIFLL	100	263	-	-	2846	26667	1
1 0573	10	POL	773	ILRGTSFYV	80	313	-	-	-	-	1

1. Frequency of entire sequence amongst isolates scanned.
2. Number of supertype alleles bound. Peptides binding 3 or more alleles are considered degenerate.
3. A dash (-) indicates IC50

Table XXVII: in vitro binding of conserved HBV-derived peptides to HLA-A3-supertype alleles.

Peptide	AA	Molecule	1st Pos	Sequence	Consv. <sup>1</sup>	A3-supertype binding capacity (IC50 nM)							Alleles bound
						A*03	A*11	A*3101	A*3301	A*6801			
26 0535	11	X NUC FUS	299	GVWIRTPAYR	95	58	35	3 0	40	12			5
1147.16	11	pol	149	HTLWKAGILYK	100	20	14	486	403	42			5
26 0539	11	POL	376	RLVVDQSFSR	95	39	2 0	7 0	24	1 0			5
26 0149	9	X	69	CALRFTSAR	85	3335	261	12	3 6	11			4
1 0993	9	X	130	KYFVLGCCR	75	262	73	30	408	2667			4
26 0153	9	X	64	SSAGPCALR	90	1375	43	55	181	11			4
1083.01	11	Core	141	STLPETTVRR	95	733	4 0	180	181	26			4
20 0130	9	pol	655	AFTFSPTYK	95	42	150	3103	13182	296			3
26 0008	8	POL	656	FTFSPTYK	95	193	136	1286	1000	73			3
1 0219	9	X	1550	FVLGGCRHK	80	169	316	1500	744	103			3
1069.20	10	POL	388	LVDQSFSR	100	6875	17	692	126	16			3
1069 16	9	POL	47	NVSLPWTHK	100	134	105	3	2900	250			3
1090.10	10	POL	665	QAFTEPTYK	95	244	11	18000	5088	6 7			3
1090.11	9	POL	531	SAICSVVR	95	1897	29	1200	446	21			3
20 0131	9	pol	524	SVVRRAPFH	95	100	10	621	-	500			3
26 0545	11	X NUC FUS	318	TLPETTVRRR	95	22000	375	2951	408	13			3
26 0023	8	X NUC FUS	296	VSEGVWIR	90	2750	207	240	1074	222			3
1142.05	9	POL	55	KVGNFTGLY	95	52	353	-	-	-			2
1142.06	9	POL	623	PVNRPIDWK	85	355	43	-	-	8889			2
1 0975	9	POL	106	RLKLMPAR	75	116	-	5 8	592	-			2
1 0562	10	POL	576	SLGHLNPNK	75	55	77	-	-	-			2
1069.21	10	NUC	170	STLPETTVR	95	15714	100	2250	1208	320			2
1069.22	10	NUC	171	TLPETTVRR	95	15714	261	-	2417	182			2
1069.15	10	POL	150	TLWKAGILYK	100	2 1	17	3529	29000	615			2
1 0215	9	X	105	TTDLAAYFK	75	18333	6 5	-	24167	471			2
1069.17	10	POL	369	VTGGVFLVDK	100	282	65	-	-	3636			2
1069.19	9	POL	389	VVDQSFSR	100	7333	80	13846	1706	242			2
26 0026	8	POL	168	ASFQGSPPY	100	239	26	-	-	20000			2
26 0549	11	ENV	389	LPIHFCLVYV	100	478	10000	2609	644	82			2
26 0550	11	POL	528	RAEPHCLAFSY	95	92	15	667	26364	2667			2
1090.04	10	POL	746	GTDNVSVLSR	90	11000	143	6000	15263	10000			1
1069.04	10	POL	149	HTLWKAGILY	100	250	7500	-	8529	6667			1
1 0205	9	POL	771	ILRGTSFVY	80	250	-	-	-	-			1
1090.08	9	NUC	148	LVSFGWIR	90	3929	500	-	-	-			1
1039.01	10	ENV	360	MMWYWGPSLY	85	220	7500	-	-	26667			1
1 0584	10	X	104	STIDLEAYFK	75	1667	2 2	-	-	-			1
1147.17	11	pol	735	GTDNVSVLSRK	90	786	11	-	-	-			1
1147.18	11	pol	357	RVTGGVFLVDK	100	578	207	-	-	-			1
1099.03	9	POL	150	TLWKAGILY	100	85	7500	-	-	-			1
1090.15	10	POL	549	YMDDVVLGAK	90	333	1395	-	-	-			1
26 0024	8	POL	50	VSLPWTHK	100	846	353	5806	22308	20000			1

1. Frequency of entire sequence amongst isolates scanned.  
2. Number of superpeptide alleles bound. Peptides binding to more than one allele are considered degenerate.  
3. A dash (-) indicates IC50

Table XXVIII: in vitro binding of conserved HBV-derived peptides to HLA-B7 supertype alleles.

Peptide	AA	Molecule	1st Pos	Sequence	Cons <sup>1</sup>	B7-supertype binding capacity (IC50 nM)					Alleles bound <sup>2</sup>
						B*0702	B*3501	B*5101	B*5301	B*5401	
1147.05	10	POL	541	EPHCLAFSYM	95	56	33	61	118	208	5
1145.04	9	ENV	324	IPRSSWAF	100	42	2.6	2.3	12	2941	4
1147.02	9	POL	440	HPAAMPHL	100	56	267	500	186	833	4
1147.06	9	X	58	LPVCAFSSA	95	115	101	500	10333	0.53	4
1147.08	9	POL	651	YPALMPLYA	95	306	150	162	664	0.63	4
988.05	9	CORE	19	LPSEDFPSV	95	1774	343	9.0	120	4.8	4
1145.08	9	POL	541	EPHCLAFSY	95	3	14	83	17	503	3
19.0014	8	POL	640	YPALMPLY	190	13750	28	13	207	1786	3
26.0570	11	pol	640	YPALMPLYACI	95	1375	-	117	291	143	3
1147.04	10	POL	365	TPARVTGGVF	90	17	72	-	939	16667	2
15.0034	9	ENV	390	LPFFCLWV	100	-	-	57	2325	53	2
20.0140	9	POL	723	LPHITAEIL	85	1375	114	1058	30	20000	2
19.0006	8	ENV	340	VPEVQWV	95	5500	-	0.29	-	91	2
19.0007	8	ENV	379	LPFFCLW	100	-	-	153	66	2857	2
19.0010	8	POL	1	MPLSYQHF	100	-	742	458	251	526	2
19.0011	8	POL	429	HPAAMPIL	100	85	18000	18	2514	625	2
19.0012	8	POL	511	SPFLAQF	95	10	8000	306	10333	1075	2
26.0566	11	pol	511	SPFLAQFSA	95	67	-	-	-	0.83	2
1147.01	9	POL	789	DPSRGRLGL	90	458	-	-	-	-	1
16.0182	10	X	67	GPCALRFSA	90	61	-	-	-	2857	1
20.0273	10	POL	440	HPAAMPILV	85	344	3600	705	664	588	1
15.0030	9	ENV	191	IPQSLDSWV	90	-	-	27500	62	-	1
15.0210	10	POL	123	LPLDKGKPY	100	-	248	27500	-	-	1
16.0006	9	ENV	25	FPDHLQDFA	90	-	8000	-	-	12	1
16.0177	10	ENV	324	IPRSSWAF	80	4231	3000	-	6643	22	1
16.0180	10	POL	644	APFTQCGYP	95	1897	-	-	-	7.1	1
16.0181	10	POL	723	LPHITAEIL	85	3056	6545	500	5813	30	1
19.0003	8	ENV	173	GPLVLVQA	95	18333	-	2895	-	1538	1
19.0005	8	ENV	313	IPRSSWA	100	13750	18000	196	-	167	1
19.0009	8	NUC	133	RPPNAPIL	100	724	-	2895	-	-	1
19.0015	8	POL	659	SPTYKAF	95	14	-	2895	-	-	1
19.0016	8	POL	769	VPSALNPA	90	5000	-	786	-	10	1
26.0554	11	pol	633	APFTQCGYPAL	95	24	7200	13750	-	1075	1
26.0559	11	pol	712	LPHITAEILAA	85	611	2667	-	775	3.6	1
26.0561	11	pol	774	NPADDSRGRL	90	458	-	-	-	-	1
26.0564	11	Core	133	RPPNAPILSTL	100	42	-	3056	-	-	1
26.0567	11	Core	49	SPHHTALRQAI	100	9.5	-	13750	18600	-	1
26.0568	11	pol	354	TPARVTGGVFL	90	58	-	-	18600	20000	1

- Frequency of entire sequence amongst isolates scanned.
- Number of supertype alleles bound. Peptides binding 3 or more alleles are considered degenerate.
- A dash (-) indicates IC50

**Table XXIX: HBV derived A1- and A24-motif containing peptides****a. A1-motif peptides**

Peptide	Molecule	Position	Sequence	Conserv.	HLA-A*0101 binding (IC50 nM)
1069.01	Core	59	LLDTASALY	75	2.1
1.0519	Core	419	DLLDTASALY	75	2.3
1069.02	pol	427	SLDVSAAFY	95	4.8
2.0239		1000	LSLDVSAAFY	95	6.0
2.0126		1521	MSTTDLEAY	75	29
1039.06	ENV	359	WMMWYWGPSLY	85	78
1090.14	pol	538	YMDDVVLGA	90	96
1090.09	pol	808	PTTGRTSLY	85	119
1069.03	pol	124	PLDKGIKPY	100	147
1069.08	env	249	ILLCLIFLL	100	192
1069.04	pol	149	HTLWKAGILY	100	381
1039.01		360	MMWYWGPSLY	85	309
1.0774	Core	416	WLWGMDIDPY	75	309
20.0254	pol	631	FAAPFTQCGY	95	368
1.0166	pol	629	KVGNFTGLY	95	368

A dash indicates IC50 nM

**b. A24 -motif peptides**

Peptide	Molecule	Position	Sequence	Conserv.	HLA-A*2402 binding (IC50 nM)
20.0271.	POL	392	SWPKFAVPNL	95	2.1
1069.23	POL	745	KYTSPFWLL	85	2.3
2.0181	POL	492	LYSHPIILGF	80	11
20.0269	ENV	236	RWMCLRRFII	95	11
20.0136	ENV	334	SWLSLLVPF	100	31
20.0137	ENV	197	SWWTSLNFL	95	32
20.0135	ENV	236	RWMCLRRFI	95	169
20.0139	POL	167	SFCGSPYSW	100	169
2.0173	POL	4	SYQHFRKLLL	75	182
2.0060		1224	GYPALMPY	95	245
13.0129	NUC	117	EYLVSFVWVI	90	353
1090.02	core	131	AYRPPNAPI	90	387
13.0073	NUC	102	WFHISCLTF	80	400
20.0138	POL	51	PWTHKVGNF	100	414

A dash indicates IC50 nM



**Table XXXa: Immunogenicity of HBV-derived A2-supermotif cross-reactive peptides**

Peptide	Sequence	Protein	XRN	Immunogenicity			overall <sup>1</sup>
				primary	transgenic	patients	
924.07	FLPSDFPSPV	HBV core 18	5	10/10	6/6	25/32 <sup>a</sup>	+
1069.06	LLVPFVQWFV	HBV env 338	5	3/4	6/9		+
1147.13	FLLAQFTSAI	HBV pol 513	5		0/3		-
1090.77	YMDDVVLGV	HBV pol 538	5		9/9		+
777.03	FLLTRLLTI	HBV env 183	4			14/23 <sup>a</sup>	+
927.15	ALMPLYACI	HBV pol 642	4	10/12	3/5	2/15 <sup>a</sup>	+
1013.01	WLSLLVPFV	HBV env 335	4	2/6	5/9	23/29 <sup>a</sup>	+
1069.05	LLAQFTSAI	HBV pol 504	4	0/4	0/5		-
1132.01	LVPFVQWFV	HBV env 339	4	0/3	0/4		-
1147.14	VLDDYQGMLPV	HBV env 259	4	4/4	6/6		+
927.41	LSSNLSWL	HBV pol 992	3	0/4	0/3		-
927.42	NLSWLSLDV	HBV pol 411	3		2/8		+
927.46	KLHL YSHPI	HBV pol 489	3	0/4	4/6		+
1069.07	FLLAQFTSA	HBV pol 503	3	1/2	0/3		+
1168.02	GLSRVVARL	HBV pol 455	3			9/13 <sup>a</sup>	+

Immunogenicity evaluation derived from primary cultures, acute patients (a-Bertoni et al, J Clin Invest 100:503, b- Rehmann et al, J. Clin. Invest 97:1655, c- Nayersina et al., J Immunol 150:4659) or transgenic mice. A positive assessment (+) is assigned when responders have been noted in one of these systems.

Table XXXb: Immunogenicity of non-crossreactive HBV A2-supermotif peptides

Peptide	Sequence	Protein	Immunogenicity				overall <sup>1</sup>
			XRN	primary	transgenic	patients	
927.11	FLSLGHL	HBV pol 562	2	15/22	12/13	9/15 <sup>a</sup>	+
927.47	HLYSHPLL	HBV pol 1076	2		10/14		+
1039.03	MMWYWGPSL	HBV env 360	2	3/4	0/4		+
1069.12	YLHTLWKAGV	HBV pol 147	2	2/4			+
1137.02	LLDYQGMLPV	HBV env 260	2	1/2	0/4		+
1142.07	GLGWSPQA	HBV env 62	2	3/4	5/6		+
1.0573	ILRGTSFVYV	HBV pol 773	1			3/7 <sup>b</sup>	+
1013.14	VLQAGFEL	HBV env 177	1	0/4	5/12		+
1069.10	LLPIFFCLWV	HBV env 378	1	3/3	0/4	2/5 <sup>c</sup>	+
1069.13	PLLPIFFCL	HBV env 377	1	0/4	7/12		+
1090.06	LLVLQAGFEL	HBV env 175	1	1/5	0/4		+
1090.12	YLVSFQVWI	HBV nuc 118	1	9/9			+
1.0518	GLSPTVWLSV	HBV env 338	1			3/9 <sup>c</sup>	+
1090.14	YMDDVVLGA	HBV pol 538	1	2/7	2/5	2/7 <sup>b</sup>	+

Immunogenicity evaluation derived from primary cultures, acute patients (a-Bertoni et al, J Clin Invest 100:503, b- Rehmann et al., J. Clin. Invest 97:1655, c- Nayersina et al., J Immunol 150:4659) or transgenic mice. A positive assessment (+) is assigned when responders have been noted in one of these systems.

Table XXXc: Cross-recognition of HBV pol 538 and a Lamivudine induced pol 538 variant by CTL induced with a pol 538 analog<sup>a</sup>.

Stimulating peptide	Day 6 CTL response (ALU)	
	HBV pol 538 (YMDDVVLGA) <sup>b</sup>	HBV pol 538 mutant (YVDDVVLGA)
HBV pol 538	27.8	54.2
HBV pol 538 mutant	35.3	27.9

a. CTLs were induced using the 1090.77 analog of HBV pol 538 (peptide 1090.14). 1090.77 was encoded in the DNA minigene pEP2.AOS.

b. Values shown represent the geometric mean of ALU from 2 independent cultures. Peptides loaded onto target cells were 1090.14 (HBV pol 538) or 1353.02 (a Lamivudine induced mutant of pol 538).

Table XXXIa: Immunogenicity of HBV-derived A3-supermotif cross-reactive peptides

Peptide	Sequence	Protein	XRN	Immunogenicity		
				primary	transgenic	overall <sup>1</sup>
1147.16	HTLWKAGIL YK	HBV POL 149	5	0/6	3/3	1/22
1083.01	STLPETTVVRR	HBV core 141	4	3/5	6/6	8/32
1150.51	GSTHVSWPK	HBV pol 398	4		3/6	
1.0219	FVLGGCRHK	HBV adr "X" 1550	3	0/4		
1069.16	NVSIPWTHK	HBV pol 47	3	0/8	0/3	1/21
1069.20	LVVDFSQFSR	HBV pol 388	3	0/4	6/6	1/22
1090.10	QAFTFSPTYK	HBV pol 665	3	3/6	0/3	3/21
1090.11	SAICSVVRR	HBV pol 531	3	1/4		2/22

1. Immunogenicity evaluation derived from primary cultures, Bertoni et al, J Clin Invest 100:503 or transgenic mice. A positive assessment (+) is assigned when responders have been noted in one of these systems. A negative assessment (-) indicates that no responders when examined.

Table XXXIb: Immunogenicity of non-crossreactive HBV A3-supermotif peptides

Peptide	Sequence	Protein	XRN	Immunogenicity		
				primary	transgenic	overall <sup>1</sup>
1069.15	TLWKAGIL YK	HBV pol 150	2	3/8	0/3	5/28
1142.05	KVGNFTGLY	HBV adr POL 629	2		0/3	2/22

1. Immunogenicity evaluation derived from primary cultures, Bertoni et al, J Clin Invest 100:503 or transgenic mice. A positive assessment (+) is assigned when responders have been noted in one of these systems. A negative assessment (-) indicates that no responders when examined.

**Table XXXIIa: Immunogenicity of HBV B7-supermotif cross-reactive peptides**

Peptide	Sequence	Protein	Immunogenicity			
			XRN	primary	transgenic	patients overall <sup>1</sup>
1147.05	FPHCLAFSYM	HBV POL 530	5	1/3	0/12	+
988.05	LPSEDFPSV	HBV core 19-27	4		2/16	+
1145.04	IPIPSSWAF	HBV ENV 313	4	0/4	1/12	+
1147.02	HPAAMPHL	HBV POL 429	4	0/5	0/12	-
1147.06	LPVCAFFSA	HBV X 58	4	1/4		+
1147.08	YPALMPLYA	HBV POL 640	4		0/12	-
1145.08	FPHCLAFSY	HBV POL 541	3	0/4		-

1. Immunogenicity evaluation derived from primary cultures, Bertoni et al, J Clin Invest 100:503 or transgenic mice. A positive assessment (+) is assigned when responders have been noted in one of these systems. A negative assessment (-) indicates that no responders when examined.

**Table XXXIIb: Immunogenicity of non-crossreactive HBV B7-supermotif peptides**

Peptide	Sequence	Protein	Immunogenicity			
			XRN	primary	transgenic	patients overall <sup>1</sup>
1147.04	TPARVTGGVF	HBV POL 354	2		2/12	+

1. Immunogenicity evaluation derived from primary cultures, Bertoni et al, J Clin Invest 100:503 or transgenic mice. A positive assessment (+) is assigned when responders have been noted in one of these systems. A negative assessment (-) indicates that no responders when examined.

Table XXXIII. Candidate HBV-derived HTL epitopes

Selection criteria	Peptide	Mol	1st Pos	Conservancy		Sequence
				Core	Total	
DR-supermotif	F107.01	ENV	249	100	95	ILLCLIFLLVLLDY
	F107.02	ENV	252	95	95	LCLIFLLVLLDYQGM
	1280.17	ENV	258	90	90	LVLLDYQGMLPVCPL
	1186.22	ENV	332	100	100	RFSWLSLLVPFVQWF
	1186.15	ENV	339	95	95	LVPFVQWFGVLSPTV
	1186.06	ENV	342	95	95	FVQWFGVLSPTVWLS
	1186.03	NUC	19	85	85	ASKLCLGWLWGMDID
	1186.12	NUC	24	85	85	LGWLWGMDIDPYKEF
	857.02	NUC	50		90	PHHTALRQAILCWGELMTLA
	1186.23	NUC	98	85	85	RQLLWFHISCLTFGR
	27.0279	NUC	117		90	EYLVSGVWIRTPPA
	27.0280	NUC	123	95	95	GVWIRTPPAYRPPNA
	1186.20	NUC	129	100	95	PPAYRPPNAPILSTL
	1186.16	NUC	136	100	95	NAPILSTLPETTIVR
	1186.01	POL	38	95	95	AEDNLGNLNVSIPW
	1186.17	POL	45	100	95	NLNVSIPWTHKVGNF
	27.0281	POL	145	100	100	RHYLHTLWKAGILYK
	1280.13	POL	406	95	95	KFAVPNLQSLTNLLS
	27.0283	POL	409		85	VPNLQSLTNLLSSNL
	F107.03	POL	412	90	90	LQSLTNLLSSNLSWL
	1186.28	POL	416	90	90	TNLLSSNLSWLSLDV
	1186.27	POL	420	100	85	SSNLSWLSLDVSAAF
	F107.04	POL	523	95	95	PFLAQFTSAICSVV
	1186.10	POL	526	95	95	LAQFTSAICSVVRA
	1186.04	POL	534	95	95	CSVVRRAPFHCLAFS
	F107.05	POL	538	95	95	RRAPFHCLAFSYMDD
	1186.02	POL	546	90	90	AFSYMDDVVLGAKSV
	1186.05	POL	629	85	85	DWKVCQRIVGLLGFA
	1280.21	POL	637	95	95	VGLLGFAAPFTQCGY
	27.0278	POL	643		95	AAPFTQCGYPALMPL
	1186.21	POL	648	95	95	QCGYPALMPLYACIQ
	1280.14	POL	694	95	95	LCQVFADATPTGWGL
	27.0282	POL	750	85	85	SVVLSRKYTSFPWLL
		X	13	95	90	RDVLCRLPVGAEARG
	1186.07	X	50	95	90	GAHLSLRGLPVCAFS
	1186.29	X	60	95	90	VCAFSSAGPCALRFT
Algorithm	1280.20	ENV	330	100	80	SVRFSWLSLLVPFVQ
	1280.19	NUC	28	85	80	RDLLDTASALYREAL
	1298.02	POL	56	90	55	VGNFTGLYSSTVPVF
	1298.03	POL	571	95	75	TNFLSLGIHLNPNK
	1298.05	POL	651	95	55	YPALMPLYACIQSKQ
	1298.06	POL	664	95	60	KQAFTFSPYKAFLC
	1280.181	POL	722	85	80	PLPIHTAELLAACFA
	1280.09	POL	774	90	80	GTSFVYVPSALNPAD
DR3-motif	795.05	ENV	10		95	PLGFFPDHQLDP
	35.0090	ENV	312	95	90	FLLVLLDYQGMLPVC
	CF-03	NUC	28	85	80	RDLLDTASALYREALESPH
	35.0091	POL	18	90	65	AGPLEEELPRLADEG
	35.0092	POL	34	100	85	NRRVAEDNLGNLNV
	35.0093	POL	96	85	60	VGPLTVNEKRRRLKI
	35.0094	POL	120	100	100	TKYLPLDKGIKPYYP

Table XXXIII. Candidate HBV-derived HTL epitopes

Selection criteria	Peptide	Mol	1st Pos	Conservancy		Sequence
				Core	Total	
DR-supermotif	F107.01	ENV	249	100	95	ILLCLIFLLVLLDY
	F107.02	ENV	252	95	95	LCLIFLLVLLDYQGM
	1280.17	ENV	258	90	90	LVLLDYQGMLPVCPL
	1186.22	ENV	332	100	100	RFSWLSLLVPFVQWF
	1186.15	ENV	339	95	95	LVPFVQWFVGLSPTV
	1186.06	ENV	342	95	95	FVQWFVGLSPTVWLS
	1186.03	NUC	19	85	85	ASKLCLGWLWGMDID
	1186.12	NUC	24	85	85	LGWLWGMDIDPYKEF
	857.02	NUC	50		90	PHHTALRQAILCWGELMTLA
	1186.23	NUC	98	85	85	RQLLWFHISCLTFGR
	27.0279	NUC	117		90	EYLVSGVWIRTPPA
	27.0280	NUC	123	95	95	GVWIRTPPAYRPPNA
	1186.20	NUC	129	100	95	PPAYRPPNAPILSTL
	1186.16	NUC	136	100	95	NAPILSTLPETTIVR
	1186.01	POL	38	95	95	AEDLNLGNLNV SIPW
	1186.17	POL	45	100	95	NLNV SIPWTHKVG NF
	27.0281	POL	145	100	100	RHYLHTLWKAGILYK
	1280.13	POL	406	95	95	KFAVPNLQSLTNLLS
	27.0283	POL	409		85	VPNLQSLTNLLSSNL
	F107.03	POL	412	90	90	LQSLTNLLSSNL SWL
	1186.28	POL	416	90	90	TNLLSSNL SWLSLDV
	1186.27	POL	420	100	85	SSNL SWLSLDV SAAF
	F107.04	POL	523	95	95	PFLLAQFTSAICSVV
	1186.10	POL	526	95	95	LAQFTSAICSVV RRA
	1186.04	POL	534	95	95	CSVVRRAPFHCLAFS
	F107.05	POL	538	95	95	RRAFPHCLAFSYMDD
	1186.02	POL	546	90	90	AFSYMDDVVLGAKSV
	1186.05	POL	629	85	85	DWKVCQRIVGLLGFA
	1280.21	POL	637	95	95	VGLLGFAAPFTQCGY
	27.0278	POL	643		95	AAPFTQCGYPALMPL
	1186.21	POL	648	95	95	QCGYPALMPLYACIQ
	1280.14	POL	694	95	95	LCQVFADATPTGWGL
	27.0282	POL	750	85	85	SVVLSRKYTSFPWLL
		X	13	95	90	RDVLCRLRPVGAESRG
	1186.07	X	50	95	90	GAHLSLRGLPVCAFS
	1186.29	X	60	95	90	VCAFSSAGPCALRFT
Algorithm	1280.20	ENV	330	100	80	SVRFSWLSLLVPFVQ
	1280.19	NUC	28	85	80	RDLLDTASALYREAL
	1298.02	POL	56	90	55	VGNFTGLYSSTVPVF
	1298.03	POL	571	95	75	TNLLSLGIHLNPNK
	1298.05	POL	651	95	55	YPALMPLYACIQSKQ
	1298.06	POL	664	95	60	KQAF TFSPTYKAF LC
	1280.181	POL	722	85	80	PLPIHTAELLAACFA
	1280.09	POL	774	90	80	GTSFVYVPSALNPAD
DR3-motif	795.05	ENV	10		95	PLGFFPDHQ LDP
	35.0090	ENV	312	95	90	FLLVLLDYQGMLPVC
	CF-03	NUC	28	85	80	RDLLDTASALYREALE SPEH
	35.0091	POL	18	90	65	AGPLEEELPRLADEG
	35.0092	POL	34	100	85	NRRVAEDLNLGNLNV
	35.0093	POL	96	85	60	VGPLTVNEKRRLKLI
	35.0094	POL	120	100	100	TKYLPLDKGIKPYYP

Table XXXIII. Candidate HBV-derived HTL epitopes

Selection criteria	Peptide	Mol	1st Pos	Conservancy		Sequence
				Core	Total	
	35.0095	POL	371	100	55	GGVFLVDKNPHNTTE
	35.0096	POL	385	100	45	ESRLVVDFSQFSRGN
	1186.18	POL	422	95	85	NLSWLSLDVSAAFYH
	35.0099	POL	666	95	55	AFTFSPTYKAFLCKQ
	35.0101	X	18	95	35	LRPVGAESRGRPVSG
Lower conservancy or miscellaneous	799.01	ENV	11	80	75	PLLVLQAGFFLLTRILTIPQ
	799.02	ENV	31	95		SLDSWWTSLNFLGGTTVCLG
	799.04	ENV	71	95	75	GYRWMCLRRFIIFLIFLLC
	1298.01	ENV	117	80	40	PQAMQWNSTTFHQTL
	1280.06	ENV	180	80	80	AGFFLLTRILTIPQS
	1280.11	ENV	245	80	80	IFLIFLLCLIFLLV
	CF-08	NUC	120		90	VSFGVWIRTPPAYRPPNAPI
	1186.25	NUC	121	95	90	SFGVWIRTPPAYRPP
	1280.15	POL	501	80	80	LHLYSHPIILGFRKI
	1298.04	POL	618	80	45	KQCFRKLPVNRPIDW
	1298.07	POL	767	80	70	AANWILRGTSFVYVP
	1298.08	POL	827	80	60	PDRVHFASPLHVAWR



Table XXXIV. HLA-DR screening panels

Screening Panel	Antigen	Alleles	Representative Assay			Phenotypic Frequencies						
			Allele	Alias	Cauc.	Blk.	Jpn.	Chn.	Hisp.	Avg.		
Primary	DR1	DRB1*0101-03	DRB1*0101	(DR1)	18.5	8.4	10.7	4.5	10.1	10.4		
	DR4	DRB1*0401-12	DRB1*0401	(DR4w4)	23.6	6.1	40.4	21.9	29.8	24.4		
	DR7	DRB1*0701-02	DRB1*0701	(DR7)	26.2	11.1	1.0	15.0	16.6	14.0		
	Panel total				59.6	24.5	49.3	38.7	51.1	44.6		
Secondary	DR2	DRB1*1501-03	DRB1*1501	(DR2w2 B1)	19.9	14.8	30.9	22.0	15.0	20.5		
	DR2	DRB5*0101	DRB5*0101	(DR2w2 B2)	-	-	-	-	-	-		
	DR9	DRB1*09011,09012	DRB1*0901	(DR9)	3.6	4.7	24.5	19.9	6.7	11.9		
	DR13	DRB1*1301-06	DRB1*1302	(DR6w19)	21.7	16.5	14.6	12.2	10.5	15.1		
	Panel total				42.0	33.9	61.0	48.9	30.5	43.2		
Tertiary	DR4	DRB1*0405	DRB1*0405	(DR4w15)	-	-	-	-	-	-		
	DR8	DRB1*0801-5	DRB1*0802	(DR8w2)	5.5	10.9	25.0	10.7	23.3	15.1		
	DR11	DRB1*1101-05	DRB1*1101	(DR5w11)	17.0	18.0	4.9	19.4	18.1	15.5		
	Panel total				22.0	27.8	29.2	29.0	39.0	29.4		
Quaternary	DR3	DRB1*0301-2	DRB1*0301	(DR3w17)	17.7	19.5	0.4	7.3	14.4	11.9		
	DR12	DRB1*1201-02	DRB1*1201	(DR5w12)	2.8	5.5	13.1	17.6	5.7	8.9		
	Panel total				20.2	24.4	13.5	24.2	19.7	20.4		

Table XXXV. HBV-derived cross-reactive HLA-DR binding peptides

Peptide	Mol	1st Pos	Conservancy		Sequence	HLA-DR binding capacity (IC50 nM)									Total DR alleles bound		
			Core	Total		DR1	DR2w2 01	DR2w2 02	DR3	DR4w4	DR4w5	DR5w11	DR6	DR7		DR8	DR9
F107 03	POL	412	90	90	LQSLTNLSSNL.SWL	2.0	21	1000	-	9.4	47	294	135	167	557	682	10
1298 06	POL	664	95	60	KQAFTEPTYKALC	9.4	38	143	-	41	173	83	175	76	408	139	10
1280 06	ENV	180	80	80	AGFELLTRLITPQS	1.1	217	1053	-	8.5	253	5.6	9.5	8.1	188	58	9
1280 09	POL	774	90	80	GTSEVYVPSALNPAD	1.4	650	400	-	118	93	426	-	93	803	221	9
1186 25	NUC	121	95	90	SGFVWIRTPPAYRPP	532	827	47	-	577	603	769	17500	1042	196	938	8
27 0280	NUC	123	95	95	GVWIRTPPAYRPPNA	14	217	2.8	-	13	67	42	-	114	92	1667	8
CF-08	NUC	120	90	90	VSEGVWIRTPPAYRPPNAPI	192		105		300		426		124			5
27 0281	POL	145	100	100	RHYLHTLWKAGILYK	17	5.4	35	-	2250	1462	42	745	61	27	174	8
1186 15	ENV	339	95	95	LVPFVQWFVGLSPTV	385	13	1429	-	300	27	53	1944	2717	74	30	7
1280 15	POL	501	80	80	LHL.YSHPIILGFRKI	227	268	500	-	66	238	488	17500	-	803	1531	7
F107 04	POL	523	95	95	PFLLAQFSAICSVV	28	337	4762	-	563	317	1667	44	325	845	1271	7
1298 04	POL	618	80	45	KQCRKLLPVNRPDIW	3.3	4136	952	-	38	45	1538	814	63	845	3000	7
1298 07	POL	767	80	70	AANWILRGTSFYVVP	54	379	3279	-	882	1520	1429	140	43	196	278	7
857 02	NUC	50	90	90	PHHTALRQAILCWGLMTLA	70	9.1	211	-	85		263	193000	676	196	2273	7

a. A dash (-) indicates IC50 nM > 20,000.

Table XXXVI. HBV-derived DR3-binding peptides

Peptide	Mol	1st Pos	Conservancy		Sequence	DR3
			Core	Total		
1280.14*	POL	694	95	95	LCQVFADATPTGWGL	67
35.0096	POL	385	100	45	ESRLVVDFSQFSRGN	115
35.0093	POL	96	85	60	VGPLTVNEKRRLKLI	136
1186.27	POL	420	100	85	SSNLSWLSLDVSAAF	200
1186.18	POL	422	95	85	NLSWLSLDVSAAFYH	231

\*tested as peptide 35.0100

Table XXXVIIa: HBV Preferred CTL Epitopes

Peptide	Sequence	Protein	HLA
924.07	FLPSDFFPSV	core 18	A2
777.03	FLLTRILT	env 183	A2
927.15	ALMPYACI	pol 642	A2
1013.01	WLSLLVPFV	env 335	A2
1090.77	YMDDVVLGV	pol 538	A2/A1
1168.02	GLSRYVARL	pol 455	A2
927.11	FLLSLGIHL	pol 562	A2
1069.10	LLPIFFCLWV	env 378	A2
1069.06	LLVPFVQWFV	env 338	A2
1147.16	HTLWKAGILYK	pol 149	A3/A1
1083.01	STLPETTVVRR	core 141	A3
1069.16	NVSIPWTHK	pol 47	A3
1069.20	LVVDFSQFSR	pol 388	A3
1090.10	QAFTFSPTYK	pol 665	A3
1090.11	SAICSVVRR	pol 531	A3
1142.05	KVGNFTGLY	pol 629	A3/A1
1147.05	FPHCLAFSYM	pol 530	B7
988.05	LPSDFFPSV	core 19	B7
1145.04	IPISSWAF	env 313	B7
1147.02	HPAAMPHELL	pol 429	B7
26.0570	YPALMPYACI	pol 640	B7
1147.04	TPARVTGGVF	pol 354	B7
1.0519	DLLDTASALY	core 419	A1
2.0239	LSLDVSAAFY	pol 1000	A1
1039.06	WMMWYWGPSLY	env 359	A1
20.0269	RWMCLRRFII	env 236	A24
20.0136	SWLSLLVPF	env 334	A24
20.0137	SWWTSLNFL	env 197	A24
13.0129	EYLVSFVWVI	core 117	A24
1090.02	AYRPPNAPI	core 131	A24
13.0073	WFHISCLTF	core 102	A24
20.0271	SWPKFAVPNL	pol 392	A24
1069.23	KYTSPFWLL	pol 745	A24
2.0181	LYSHPIILGF	pol 492	A24

Table XXXVIIIb: HBV Preferred HTL epitopes

Selection Criteria	Peptide	Mol	1st Pos	Conservancy		Sequence
				Core	Total	
DR supermotif	F107.03	POL	412	90	90	LQSLTNLSSNL.SWL
	1298.06	POL	664	95	60	KQAFTEFPTYKAF.LC
	1280.06	ENV	180	80	80	AGFELLTRIL.TIPQS
	1280.09	POL	774	90	80	GTSFVYVPSALNPAD
	CF-08	CORE	120		90	VSFGVWIRTPPAYRPPNAPI
	27.0281	POL	145	100	100	RHYLHTLWKAGIL.YK
	1186.15	ENV	339	95	95	LVPFVQWFVGLSP.TV
	1280.15	POL	501	80	80	LHL.YSHPIIL.GFRKI
	F107.04	POL	523	95	95	PFLLAQFTSAICSVV
	1298.04	POL	618	80	45	KQCFRKL.PVNRPIDW
	1298.07	POL	767	80	70	AANWILRGTSFVYVP
	857.02	CORE	50		90	PHHTALRQAILCWGEL.MTLA
DR3 motif	1280.14	POL	694	95	95	LCQVFADATPTG.WGL
	35.0096	POL	385	100	45	ESRLVVD.FSQSRGN
	35.0093	POL	96	85	60	VGPLTVNEKRR.LKLI
	1186.27	POL	420	100	85	SSNLSWL.SLDVSAAF

**Table XXXVIII. Estimated population coverage by a panel of HBV derived HTL epitopes**

Antigen	Alleles	Representative assay	No. of epitopes <sup>2</sup>	Population coverage (phenotypic frequency)					
				Cauc.	Blk.	Jpn.	Chn.	Hisp.	Avg.
DR1	DRB1*0101-03	DR1	12	18.5	8.4	10.7	4.5	10.1	10.4
DR2	DRB1*1501-03	DR2w2 B1	11	19.9	14.8	30.9	22.0	15.0	20.5
DR2	DRB5*0101	DR2w2 B2	8	-	-	-	-	-	-
DR3	DRB1*0301-2	DR3	4	17.7	19.5	0.40	7.3	14.4	11.9
DR4	DRB1*0401-12	DR4w4	11	23.6	6.1	40.4	21.9	29.8	24.4
DR4	DRB1*0401-12	DR4w15	9	-	-	-	-	-	-
DR7	DRB1*0701-02	DR7	9	26.2	11.1	1.0	15.0	16.6	14.0
DR8	DRB1*0801-5	DR8w2	7	5.5	10.9	25.0	10.7	23.3	15.1
DR9	DRB1*09011,09012	DR9	10	3.6	4.7	24.5	19.9	6.7	11.9
DR11	DRB1*1101-05	DR5w11	11	17.0	18.0	4.9	19.4	18.1	15.5
DR13	DRB1*1301-06	DR6w19	7	21.7	16.5	14.6	12.2	10.5	15.1
Total <sup>1</sup>				98.5	95.1	97.1	91.3	94.3	95.1

1. Total opulation coverage has been adjusted to account for the presence of DRX in many ethnic populations. It has been assumed that the range of specificities represented by DRX alleles will mirror those of previously characterized HLA-DR alleles. The proportion of DRX incorporated under each motif is representative of the frequency of the motif in the remainder of the population. Total coverage has not been adjusted to account for unknown gene types.

2. Number of epitopes represents a minimal estimate, considering only the epitopes shown in Table 12. Additional alleles possibly bound by nested epitopes have not been accounted.

WHAT IS CLAIMED IS:

1. A peptide composition of less than 250 amino acid residues comprising a peptide epitope useful for inducing an immune response against hepatitis B virus (HBV) said epitope (a) having an amino acid sequence of about 8 to about 13 amino acid residues that have at least 65% identity with a native amino acid sequence of HBV and, (b) binding to at least one HLA class I HLA allele with an  $IC_{50}$  of less than about 500 nM.

2. The composition of claim 1, further wherein said peptide has at least 77% identity with a native HBV amino acid sequence.

3. The composition of claim 1, further wherein said peptide has 100% identity with a native HBV amino acid sequence.

4. A pharmaceutical composition comprising a peptide and a pharmaceutical carrier, wherein the peptide is a peptide of Table VII (A1 supermotif), Table VIII (A2 supermotif/A2.1 motif), Table IX (A3 supermotif), Table X (A24 supermotif), Table XI (B7 supermotif), Table XII (B27 supermotif), Table XIII (B58 supermotif), Table XIV (B62 supermotif), Table XV (A1 motif), Table XVI (A3 motif), Table XVII (A11 motif), or Table XVIII (A24 motif) comprising an  $IC_{50}$  of less than about 500 nM for at least one HLA class I molecule.

5. The pharmaceutical composition of claim 4 wherein the composition comprises the peptide in a form of nucleic acids that encode the peptide.

6. The pharmaceutical composition of claim 5 wherein the composition comprises the peptide in a form of nucleic acids that encode the epitope and one or more additional peptide(s).

7. The composition of claim 4, wherein the peptide is comprised by a longer peptide, with a *proviso* that the longer peptide is not an entire native antigen.

8. The pharmaceutical composition of claim 4 wherein the peptide is in a human dose form, and the carrier is in a human unit dose.

9. A peptide composition of claim 1 comprising an analog of a peptide epitope, wherein the peptide epitope is an epitope of Table VII (A1 supermotif), Table VIII (A2 supermotif/A2.1 motif), Table IX (A3 supermotif), Table X (A24 supermotif), Table XI (B7 supermotif), Table XII (B27 supermotif), Table XIII (B58 supermotif), Table XIV (B62 supermotif), Table XV (A1 motif), Table XVI (A3 motif), Table XVII (A11 motif), or Table XVIII (A24 motif), said analog comprising a preferred or less preferred amino acid of Table II substituted in for a starting residue, or having a deleterious residue of Table II substituted out of the starting sequence and replaced by a non-deleterious residue.

10. A peptide composition of claim 1 comprising a peptide of Table XXII.

11. A method for inducing a cytotoxic T lymphocyte response, said method comprising steps of:

providing a peptide that comprises an  $IC_{50}$  of less than about 500 nM for an HLA class I molecule, wherein the peptide is a peptide of Table VII (A1 supermotif), Table VIII (A2 supermotif/A2.1 motif), Table IX (A3 supermotif), Table X (A24 supermotif), Table XI (B7 supermotif), Table XII (B27 supermotif), Table XIII (B58 supermotif), Table XIV (B62 supermotif), Table XV (A1 motif), Table XVI (A3 motif), Table XVII (A11 motif), or Table XVIII (A24 motif); and,

administering said peptide to a human.

12. The method of claim 11, wherein the providing step provides the peptide in a form of nucleic acids that encode the peptide.

13. The method of claim 12, wherein the providing step provides the peptide in a form of nucleic acids that encode the peptide and at least one additional peptide, with a *proviso* that an additional peptide is not an entire native antigen.



14. The method of claim 11, wherein the providing step provides the peptide comprised by a longer peptide, with a *proviso* that the longer peptide is not an entire native antigen.

5 15. A method for inducing a cytotoxic T lymphocyte response, said method comprising steps of:

providing a peptide that induces a cytotoxic T cell response *in vitro* and/or *in vivo*, wherein the peptide is a peptide of Table VII (A1 supermotif), Table VIII (A2 supermotif/A2.1 motif), Table IX (A3 supermotif), Table X (A24 supermotif), Table XI  
10 (B7 supermotif), Table XII (B27 supermotif), Table XIII (B58 supermotif), Table XIV (B62 supermotif), Table XV (A1 motif), Table XVI (A3 motif), Table XVII (A11 motif), Table XVIII (A24 motif) or Table XXIII; and,

administering said pharmaceutical composition to a human.

15 16. The method of claim 15, wherein the providing step provides the peptide in a form of nucleic acids that encode the peptide.

17. The method of claim 16, wherein the providing step provides the peptide in a form of nucleic acids that encode the peptide and at least one additional  
20 peptide, with a *proviso* that an additional peptide is not an entire native antigen.

18. The method of claim 15, wherein the providing step provides the peptide comprised by a longer peptide, with a *proviso* that the longer peptide is not an entire native antigen.

25 19. The method of claim 15, wherein the providing step comprises a peptide that induces a cytotoxic T cell response when complexed with an HLA class I molecule and is presented to an HLA class I-restricted cytotoxic T cell.

30 20. A peptide composition of less than 250 amino acid residues comprising a peptide epitope useful for inducing an immune response against hepatitis B virus (HBV) said epitope (a) having an amino acid sequence of about 6 to about 25 amino acid residues that have at least 65% identity with a native amino acid sequence of HBV

and, (b) binding to at least one HLA class II HLA allele with an IC<sub>50</sub> of less than about 1000 nM.

21. The composition of claim 20, further wherein said peptide has at  
5 least 77% identity with a native HBV amino acid sequence.

22. The composition of claim 20, further wherein said peptide has  
100% identity with a native HBV amino acid sequence.

10 23. A pharmaceutical composition comprising:  
a human dose form of a peptide of Table XIX or Table XX that comprises  
an IC<sub>50</sub> of less than about 1,000 nM for at least one HLA DR molecule of an HLA DR  
supertype; and,  
a human dose of a pharmaceutically acceptable carrier.

15 24. The pharmaceutical composition of claim 23 wherein the  
composition comprises the peptide in a form of nucleic acids that encode the peptide.

20 25. The pharmaceutical composition of claim 24 wherein the  
composition comprises the peptide in a form of nucleic acids that encode the peptide and  
at least one additional peptide, with a *proviso* that an additional peptide is not an entire  
native antigen.

25 26. The composition of claim 25, wherein the peptide is comprised by  
a longer peptide, with a *proviso* that the longer peptide is not an entire native antigen.

27. A peptide composition of claim 20 comprising an analog of a  
peptide epitope of Table XIX or Table XX, said analog comprising a preferred or less  
preferred amino acid of Table III substituted in for a starting residue, and/or having a  
30 deleterious residue of Table III substituted out of the starting sequence and replaced by a  
non-deleterious residue.

28. A method for inducing a helper T lymphocyte response, said method comprising steps of:

providing a peptide that comprises an  $IC_{50}$  of less than about 1,000 nM for an HLA class II molecule, wherein the peptide is a peptide of Table XIX or Table XX;

and,

administering said peptide to a human.

29. The method of claim 28, wherein the providing step provides the peptide in a form of nucleic acids that encode the peptide.

30. The method of claim 29, wherein the providing step provides the peptide in a form of nucleic acids that encode the peptide and at least one additional peptide, with a *proviso* that an additional peptide is not an entire native antigen.

31. The method of claim 28, wherein the providing step provides the peptide comprised by a longer peptide, with a *proviso* that the longer peptide is not an entire native antigen.

32. A method for inducing a helper T lymphocyte response, said method comprising steps of:  
providing a peptide that induces a helper T cell response *in vitro* and/or *in vivo*, wherein the peptide is a peptide of Table XIX or Table XX; and,  
administering said pharmaceutical composition to a human.

33. The method of claim 32, wherein the providing step provides the peptide in a form of nucleic acids that encode the peptide.

34. The method of claim 33, wherein the providing step provides the peptide in a form of nucleic acids that encode the peptide and at least one additional peptide, with a *proviso* that an additional peptide is not an entire native antigen.

35. The method of claim 32, wherein the providing step provides the peptide comprised by a longer peptide, with a *proviso* that the longer peptide is not an entire native antigen.

36. The method of claim 32, wherein the providing step comprises a peptide that induces a helper T cell response when complexed with an HLA class II molecule and is presented to an HLA class I-restricted helper T cell.

37. A vaccine for preventing or treating HBV infection that induces a protective or therapeutic immune response, wherein said vaccine comprises:

at least one peptide selected from Table(s) VII-XX or Table XXII; and,  
a pharmaceutically acceptable carrier.

38. A kit for a vaccine that induces a protective or therapeutic immune response to HBV, said vaccine comprising:

at least one peptide selected from Table(s) VII-XX or Table XXII;  
a pharmaceutically acceptable carrier; and,  
instructions for administration to a patient.

39. A method for monitoring or evaluating an immune response to HBV or an epitope thereof in a patient having a known HLA type, the method comprising:

incubating a T lymphocyte sample from the patient with a peptide selected from Table(s) VII-XX or Table XXII, wherein that peptide bears a motif corresponding to at least one HLA allele present in said patient; and,  
detecting the presence of a T lymphocyte that recognizes the peptide.

40. The method of claim 39, wherein the peptide is comprised by a tetrameric complex.

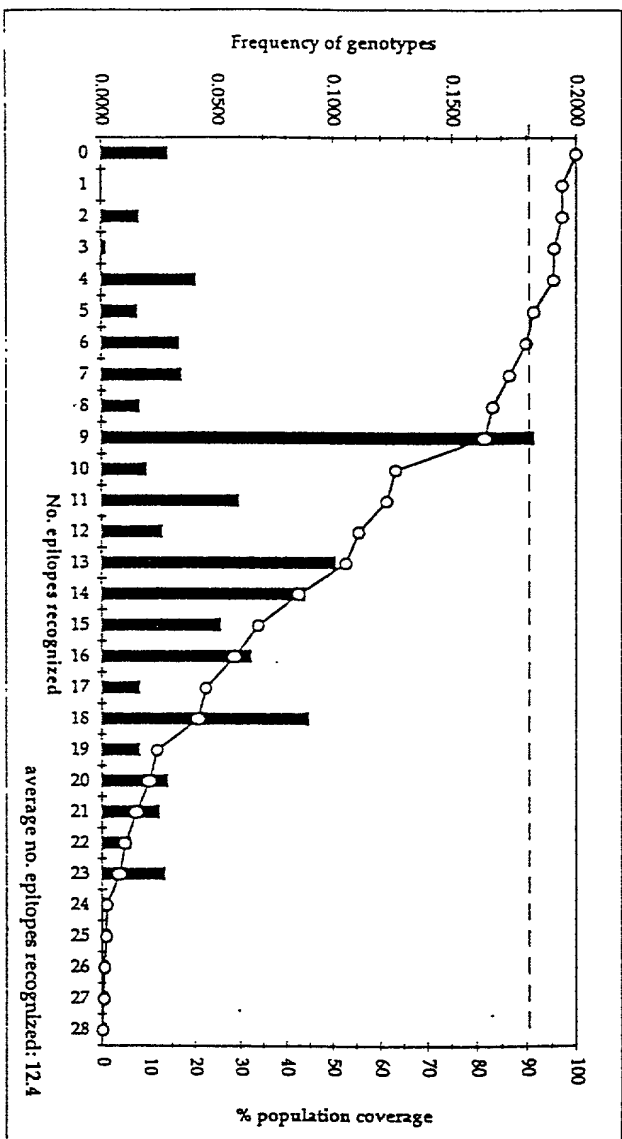
## ABSTRACT OF THE DISCLOSURE

This invention uses our knowledge of the mechanisms by which antigen is recognized by T cells to develop epitope-based vaccines directed towards HBV. More specifically, this application communicates our discovery of pharmaceutical compositions and methods of use in the prevention and treatment of HBV infection.

SF 203969 v2

10

Figure 1. Monte Carlo population coverage analysis for  
HBV candidate epitopes



#### Monte Carlo population coverage analysis for HBV candidate epitopes

Plot of total frequency of genotypes as a function of the number of HBV candidate epitopes bound by HLA-A and B alleles, in an average population. Genotype values were derived by averaging the gene frequencies in Caucasian, North American Black, Japanese, Chinese, and Hispanic populations. Also shown is the cumulative frequency of genotypes.

Using currently available HLA typing data, a residual fraction (about 15%) of the genes, in an average population, are unspecified. To arrive at 100% accounting of genes, a fraction of the residual has been added for each HLA population cluster in proportion to the relative frequency of the cluster within the HLA specified population.

A. pMin.1-No PADRE

PADRE deleted										
sig seq	HBV	HBV	HIV	HBV	HBV	HIV	HBV	HIV	HBV	
	Pol 149	Core 18	Env 120	Pol 551-V	Pol 455	Pol 476	Core 141	Env 49	Env 335	

B. pMin.1-Anchor

Pol 538 native anchor (A at P9)										
sig seq	HBV		HBV	HIV	HBV	HBV	HIV	HBV	HIV	HBV
	Pol 149	PADRE	Core 18	Env 120	Pol 551-A	Pol 455	Pol 476	Core 141	Env 49	Env 335

C. pMin.1-No Sig

Signal sequence deleted										
sig seq	HBV		HBV	HIV	HBV	HBV	HIV	HBV	HIV	HBV
	Pol 149	PADRE	Core 18	Env 120	Pol 551-V	Pol 455	Pol 476	Core 141	Env 49	Env 335

D. pMin.1-Switch

Position of HBV Env 335 and HBV Pol 455 switched										
sig seq	HBV		HBV	HIV	HBV	HBV	HIV	HBV	HIV	HBV
	Pol 149	PADRE	Core 18	Env 120	Pol 551-V	Env 335	Pol 476	Core 141	Env 49	Pol 455

FIGURE 2

## DECLARATION

As a below named inventor, I declare that:

My residence, post office address and citizenship are as stated below next to my name; I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural inventors are named below) of the subject matter which is claimed and for which a patent is sought on the invention entitled: **INDUCING CELLULAR IMMUNE RESPONSES TO HEPATITIS B VIRUS USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS** the specification of which \_\_\_\_\_ is attached hereto or X was filed on July 8, 1999 as Application No. \_\_\_\_\_ and was amended on \_\_\_\_\_ (if applicable).

I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above. I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56. I claim foreign priority benefits under Title 35, United States Code, Section 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed.

### Prior Foreign Application(s)

Country	Application No.	Date of Filing	Priority Claimed Under 35 USC 119

I hereby claim the benefit under Title 35, United States Code § 119(e) of any United States provisional application(s) listed below:

Application No.	Filing Date
60/013,363	March 13, 1996

I claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, Section 1.56 which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

Application No.	Date of Filing	Status
08/820,360	March 12, 1997	Pending
09/189,702	November 10, 1998	
08/205,713	March 4, 1994	
08/159,184	November 29, 1993	
08/073,205	June 4, 1993	
08/027,146	March 5, 1993	



Full Name of Inventor 1:	Last Name: <b>SETTE</b>	First Name: <b>ALESSANDRO</b>	Middle Name or Initial:	
Residence & Citizenship:	City: <b>La Jolla</b>	State/Foreign Country: <b>California</b>	Country of Citizenship: <b>Italy</b>	
Post Office Address:	Post Office Address: <b>5551 Linda Rosa Avenue</b>	City: <b>La Jolla</b>	State/Country: <b>California</b>	Postal Code: <b>92037</b>
Full Name of Inventor 2:	Last Name: <b>SIDNEY</b>	First Name: <b>JOHN</b>	Middle Name or Initial:	
Residence & Citizenship:	City: <b>La Jolla</b>	State/Foreign Country: <b>California</b>	Country of Citizenship: <b>United States</b>	
Post Office Address:	Post Office Address: <b>4218 Corte de la Siena</b>	City: <b>La Jolla</b>	State/Country: <b>California</b>	Postal Code: <b>92130</b>
Full Name of Inventor 3:	Last Name: <b>SOUTHWOOD</b>	First Name: <b>SCOTT</b>	Middle Name or Initial:	
Residence & Citizenship:	City: <b>Santee</b>	State/Foreign Country: <b>California</b>	Country of Citizenship: <b>United States</b>	
Post Office Address:	Post Office Address: <b>10679 Strathmore Dr.</b>	City: <b>Santee</b>	State/Country: <b>California</b>	Postal Code: <b>92071</b>
Full Name of Inventor 4:	Last Name: <b>VITIELLO</b>	First Name: <b>MARIA</b>	Middle Name or Initial: <b>A.</b>	
Residence & Citizenship:	City: <b>La Jolla</b>	State/Foreign Country: <b>California</b>	Country of Citizenship: <b>Italy</b>	
Post Office Address:	Post Office Address: <b>7522 High Avenue</b>	City: <b>La Jolla</b>	State/Country: <b>California</b>	Postal Code: <b>92037</b>
Full Name of Inventor 5:	Last Name: <b>LIVINGSTON</b>	First Name: <b>BRIAN</b>	Middle Name or Initial: <b>D.</b>	
Residence & Citizenship:	City: <b>San Diego</b>	State/Foreign Country: <b>California</b>	Country of Citizenship: <b>United States</b>	
Post Office Address:	Post Office Address: <b>13555 Chaco Court</b>	City: <b>San Diego</b>	State/Country: <b>California</b>	Postal Code: <b>92129</b>
Full Name of Inventor 6:	Last Name: <b>CELIS</b>	First Name: <b>ESTEBAN</b>	Middle Name or Initial:	
Residence & Citizenship:	City: <b>Rochester</b>	State/Foreign Country: <b>Minnesota</b>	Country of Citizenship: <b>United States</b>	
Post Office Address:	Post Office Address: <b>3683 Wright Road S.W.</b>	City: <b>Rochester</b>	State/Country: <b>Minnesota</b>	Postal Code: <b>55902</b>
Full Name of Inventor 7:	Last Name: <b>KUBO</b>	First Name: <b>RALPH</b>	Middle Name or Initial: <b>T.</b>	
Residence & Citizenship:	City: <b>Carlsbad</b>	State/Foreign Country: <b>California</b>	Country of Citizenship: <b>United States</b>	
Post Office Address:	Post Office Address: <b>6921 Pear Tree Drive</b>	City: <b>Carlsbad</b>	State/Country: <b>California</b>	Postal Code: <b>92009</b>
Full Name of Inventor 8:	Last Name: <b>GREY</b>	First Name: <b>HOWARD</b>	Middle Name or Initial: <b>M.</b>	
Residence & Citizenship:	City: <b>La Jolla</b>	State/Foreign Country: <b>California</b>	Country of Citizenship: <b>United States</b>	
Post Office Address:	Post Office Address: <b>1461 Caminito Batea</b>	City: <b>La Jolla</b>	State/Country: <b>California</b>	Postal Code: <b>92037</b>
Full Name of Inventor 9:	Last Name: <b>CHESNUT</b>	First Name: <b>ROBERT</b>	Middle Name or Initial: <b>W.</b>	
Residence & Citizenship:	City: <b>Cardiff-by-the-Sea</b>	State/Foreign Country: <b>California</b>	Country of Citizenship: <b>United States</b>	
Post Office Address:	Post Office Address: <b>1473 Kings Cross Drive</b>	City: <b>Cardiff-by-the-Sea</b>	State/Country: <b>California</b>	Postal Code: <b>92007</b>

I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Signature of Inventor 1	Signature of Inventor 2	Signature of Inventor 3
<u>ALESSANDRO SETTE</u>	<u>JOHN SIDNEY</u>	<u>SCOTT SOUTHWOOD</u>
Date	Date	Date
Signature of Inventor 4	Signature of Inventor 5	Signature of Inventor 6
<u>MARIA A. VITIELLO</u>	<u>BRIAN D. LIVINGSTON</u>	<u>ESTEBAN CELIS</u>
Date	Date	Date
Signature of Inventor 7	Signature of Inventor 8	Signature of Inventor 9
<u>RALPH T. KUBO</u>	<u>HOWARD M. GREY</u>	<u>ROBERT W. CHESNUT</u>
Date	Date	Date

SP 1009533 v1